

# *Rationalising framing effects: at least one task for empirically informed philosophy*

Article

Published Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Open Access

Fisher, S. A. ORCID: <https://orcid.org/0000-0003-1115-6134>  
(2020) Rationalising framing effects: at least one task for empirically informed philosophy. *Crítica, Revista Hispanoamericana de Filosofía*, 52 (156). pp. 5-30. ISSN 1870-4905 doi: 10.22201/iifs.18704905e.2020.1221 Available at <https://centaur.reading.ac.uk/92548/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Identification Number/DOI: 10.22201/iifs.18704905e.2020.1221  
<<https://doi.org/10.22201/iifs.18704905e.2020.1221>>

Publisher: Instituto de Investigaciones Filosóficas

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## RATIONALISING FRAMING EFFECTS: AT LEAST ONE TASK FOR EMPIRICALLY INFORMED PHILOSOPHY

SARAH A. FISHER  
University of Reading  
School of Humanities  
Department of Philosophy  
sarah.fisher@univie.ac.at

**SUMMARY:** Human judgements are affected by the words in which information is presented —or ‘framed’. According to the standard gloss, ‘framing effects’ reveal counter-normative reasoning, unduly affected by positive/negative language. One challenge to this view suggests that number expressions in alternative framing conditions are interpreted as denoting lower-bounded (minimum) quantities. However, it is unclear whether the resulting explanation is a rationalising one. I argue that a number expression should only be interpreted lower-boundedly if this is what it actually means. I survey how number expressions might be assigned lower-bounded meanings, due to their conventional semantics or pragmatic enrichment in context. I argue that deciding between these possibilities requires foundational philosophical input.

**KEY WORDS:** cognitive bias, rationality, number expressions, semantics, pragmatics

**RESUMEN:** Los juicios humanos se ven afectados por las palabras en las que la información se presenta “enmarcada”. Según la glosa estándar, el “efecto marco” revela un razonamiento contra-normativo, indebidamente afectado por el lenguaje positivo/negativo. Un desafío a este punto de vista sugiere que las expresiones numéricas en condiciones marco alternativas se interpretan denotando cantidades de límite inferior (mínimas). Sin embargo, no está claro si la explicación resultante es racionalizadora. Sostengo que una expresión numérica solo debe interpretarse con límites inferiores si esto es lo que realmente significa. Examinó cómo a las expresiones numéricas se les pueden asignar significados de límite inferior, debido a su semántica convencional o al enriquecimiento pragmático en el contexto. Sostengo que decidir entre estas posibilidades requiere una aportación filosófica fundamental.

**PALABRAS CLAVE:** sesgo cognitivo, racionalidad, expresiones numéricas, semántica, pragmática

### 1. *Introduction*

Consider the following experimental scenario, which was first used in an empirical study conducted by the psychologists Amos Tversky and Daniel Kahneman (1981):

Imagine the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume the exact scientific estimates of the consequences are as follows:

In one condition, the options are as follows:

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

In another condition, the options are:

If Program C is adopted, 400 people will die.

If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.

In both conditions, participants are then asked:

Which of the two programs would you favor?

Tversky and Kahneman found that most participants in the ‘be saved’ condition favoured Program A —the ‘sure option’ (72% chose this option, compared with just 28% who favoured the ‘risky option’, Program B). In contrast, in the ‘die’ condition, most participants favoured the risky option, Program D (78%, compared with 22% who favoured the sure option, Program C). This is puzzling on the standard interpretation of the two conditions, according to which they describe exactly the same pair of prospects. The mere use of positive or negative language (‘be saved’ vs. ‘die’) seems to produce a reversal in people’s preferences —and in their attitude to risk-taking.

Tversky and Kahneman’s ‘Asian Disease Problem’ (ADP) has inspired a programme of psychological research into the effects of framing information in different words —typically, by using language that is either positively or negatively valenced. In this paper, I will focus solely on ADP-style ‘risky-choice’ framing paradigms, which elicit

choices between outcomes occurring with varying degrees of probability.<sup>1</sup> In particular, I will address the question of why seemingly equivalent options end up being evaluated differently, and whether this can be a *rational* response on the part of a decision-maker. The effects of framing on choice have standardly been interpreted as evidence of counter-normative (irrational or sub-rational) cognitive processing. As Bermúdez has recently put it: “The consensus view is that susceptibility to framing effects is a paradigm of irrationality” (2018, p. 180). The idea that people are subject to such cognitive biases has proved extremely influential, both within and beyond academia, as part of a wider focus on insights from the behavioural sciences.<sup>2</sup> Over the last couple of decades, however, various challenges to the prevailing view have begun to gain momentum. Here I will focus on the idea, tested empirically by Mandel (2014), that the number expressions in framing problems may be interpreted *lower-boundedly*, as denoting *minimum* rather than *exact* quantities. I will argue that this hypothesis may support a *rationalising* explanation of framing effects, but only on the basis of a fully worked-out theory of the meanings of number expressions. Crucially, providing that foundation will, in turn, require engagement with a conceptual distinction heavily discussed in the philosophy of language, concerning the difference between semantics and pragmatics. Thus, I aim to show how empirical and philosophical research is importantly intertwined, at least in attempts to explain framing effects.

The structure of the paper is as follows: in section 2, I describe the standard gloss of risky-choice framing effects, which are thought to reveal counter-normative reasoning. In section 3, I introduce the hypothesis that recipients of the ADP interpret number expressions as denoting lower-bounded quantities. I assess the empirical support for this hypothesis and argue that, although it may provide a partial explanation of risky-choice framing effects, it does not yet support a *rationalising* explanation; that would require showing that the number expressions have lower-bounded *meanings* in the relevant context. In section 4, I survey various accounts of the semantics and pragmatics of number expressions and describe how these could

<sup>1</sup> For surveys of the relevant psychological literature, see Kühberger 1998; Kühberger, Schulte-Mecklenbeck, and Perner 1999; Levin, Schneider, and Gaeth 1998; Steiger and Kühberger 2018.

<sup>2</sup> For example, in 2010 the UK Government established the Behavioural Insights Team (also known as the ‘Nudge Unit’), which has subsequently expanded into a global social purpose company, working across public, private, and charitable sectors.

potentially license lower-bounded interpretations of the number expressions in the ADP. In section 5, I argue that endorsing any one of these accounts will ultimately depend on philosophical arguments about how to draw the boundary between semantics and pragmatics. I conclude that empirical and philosophical research remain inextricably linked in debates about framing, rationality, representation, and meaning.

## 2. *Framing Effects and (Ir)rationality*

Framing effects are standardly thought to be counter-normative because they violate the principle of ‘description invariance’, which is taken to be an essential condition for rational decision-making. Roughly speaking, the principle of description invariance requires that, when two expressions obviously describe the same possible outcome, that outcome should be represented and evaluated in the same way under each description.<sup>3</sup> It is thought to be obvious that, in each framing condition of the ‘Asian Disease Problem’ (ADP), the same pair of possible outcomes is being described and, therefore, that these should be represented and evaluated by recipients in exactly the same way. As we have seen, though, the empirical data show that people actually evaluate them differently. In Tversky and Kahneman’s experiment, a majority preferred the sure option (Program A) in the ‘be saved’ condition, whereas, in the ‘die’ condition, a majority preferred the risky option (Program D).<sup>4</sup>

Tversky and Kahneman explain the observed framing effect via their theory of decision making under risk, known as ‘Prospect Theory’ (1979). Prospect Theory is put forward as a descriptively accurate account of human decision-making, which diverges in important ways from the normative ideal embodied in classical ‘Expected Utility Theory’. I will focus solely on the following point of divergence here: whereas Expected Utility Theory requires rational agents to

<sup>3</sup> Spelling out the description invariance principle more precisely has proved difficult: for some further discussion, see Bermúdez 2009, 2018. The rough formulation provided here will suffice for present purposes.

<sup>4</sup> Not all subsequent studies have replicated this wholesale *reversal* of preferences. However, there is a very robust tendency for responses to *shift*, i.e. larger proportions of participants typically prefer the sure option in the positively-worded condition than in the negatively-worded condition (and, correspondingly, larger proportions prefer the risky option in the negatively-worded condition than in the positively worded condition). For further discussion of this point, see Kühberger 1998, p. 30; Levin et al. 1998, p. 153.

have consistent attitudes to risk, Prospect Theory predicts that people are relatively *risk-averse* when faced with potential *gains*, and *risk-seeking* when faced with potential *losses*. To see the distinction, consider a choice between the following options:

- (1) A sure gain of \$200.
- (2) A 50% chance of gaining \$400 and a 50% chance of gaining \$0.

Compare this with the choice between the following options:

- (3) A sure loss of \$200.
- (4) A 50% chance of losing \$400 and a 50% chance of losing \$0.

According to Expected Utility Theory, a rational agent who is risk-averse will select (1) and (3), while a rational agent who is risk-seeking will select (2) and (4). However, that theory rules out the alternative combinations —(1) and (4), or (2) and (3)— since these would imply a change in risk attitude, based purely on whether one stands to *gain* or *lose* the \$200.

Prospect Theory predicts, in contrast, the selection of (1) and (4). This is because people are assumed to be risk-averse when facing gains (as in the choice between (1) and (2)) but risk-seeking when facing losses (as in the choice between (3) and (4)). The shift in risk attitude is dubbed a ‘reflection effect’ (since it seems to be caused by ‘reflecting’ the prospects around the zero-utility point).

The core idea is further extended to tasks in which the prospects merely *seem* like gains or losses. As Kahneman and Tversky put it:

The reference point usually corresponds to the current asset position, in which case gains and losses coincide with the actual amounts that are received or paid. However, the location of the reference point, and the consequent coding of outcomes as gains or losses, can be affected by the formulation of the offered prospects, and by the expectations of the decision maker. (p. 273)

A prominent way of making prospects seem like gains or losses is by describing them with positive or negative wording, as in the two conditions of the ADP. In the first condition, the options are said to *sound like gains* because of the focus on the people who will ‘be saved’. Therefore, participants tend to be *risk-averse* in this condition, preferring the sure option (Program A). Conversely, it is held

that the options in the second condition *sound like losses* because of the focus on the people who will ‘die’. And this is argued to explain why participants tend to be *risk-seeking* in the ‘die’ condition, preferring the risky option (Program D).

Offering an explanation of the underlying psychological mechanism at work, Kahneman and Tversky (1979, pp. 273–274) posit an initial ‘editing’ phase in decision-making. This involves decision-makers mentally representing the choice options in particular ways. Of most relevance for current purposes, it is hypothesised that when the options are worded *positively* they are represented (or ‘coded’) as *gains*. Conversely, when the options are worded *negatively* they are coded as losses. The implications of this editing process become evident in the subsequent ‘evaluation’ phase, in which decision-makers assess the choice options before them. Where the alternatives have been coded as gains, decision-makers are predicted to make *risk-averse* choices. Conversely, where they have been coded as losses, decision-makers are predicted to make *risk-seeking* choices. As a result, the choices are held to be *doubly* counter-normative: reasoning errors are committed at both the editing phase (where outcomes are irrationally represented in different ways) and the evaluation phase (where this difference in representation irrationally affects choice). Ultimately, what is distinctive about framing effects (as opposed to mere ‘reflection effects’) is that the positive or negative wording in the choice options cues the audience to represent the outcomes as gains or losses during the *editing* phase.

It is worth noting that there is limited discussion in the literature of exactly what counts as positive or negative wording.<sup>5</sup> For my purposes here, though, it will suffice to think of the valence of a linguistic expression as depending straightforwardly on the goodness or badness of what is denoted. Thus, for example, the expression ‘be saved’ is positive because *being saved* is typically a good thing, whereas ‘die’ is negative because *dying* is typically bad. Of course, there will not always be consensus about what counts as a good or bad thing. However, I take it that the ADP scenario, which will be my primary focus here, is sufficiently clear-cut.

As noted previously, it is standardly assumed that, in each condition of the ADP, exactly the same two prospects are described. Specifically, both of the sure options are thought to describe a cer-

<sup>5</sup> For further discussion, which disentangles the effects of using positive or negative ‘root words’ from their affirmative or negational use, see Mandel 2001 and Tombu and Mandel 2015.



tain prospect of 200 people being saved and 400 dying. Meanwhile, both of the risky options are thought to describe a one-third probability that 600 people will be saved (with none dying), together with a two-thirds probability that 600 people will die (with none being saved). This is taken to be obvious, at least once the conditions are considered side by side.<sup>6</sup> The only difference, then, is that, in the first condition, the alternatives are framed positively (in terms of the numbers of people who will ‘be saved’), whereas in the second condition they are framed negatively (in terms of the numbers of people who will ‘die’). Since this is seen as a superficial linguistic difference, distinct representation of the prospects (as gains or losses) is taken to violate the normative principle of description invariance. In sum, it is standardly assumed that the pairs of alternative ‘frames’ —(5) and (7) below; and (6) and (8)— have the same *meaning*, even though they are typically *represented* differently in each framing condition.

- (5) If Program A is adopted, 200 people will be saved.
- (6) If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.
- (7) If Program C is adopted, 400 people will die.
- (8) If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.

If this is the right gloss of what is happening, framing effects clearly evidence counter-normative behaviour. However, there are various ways of challenging the claim that alternative frames have the same meaning.<sup>7</sup> Here I will focus on the ‘lower-bounding hypothesis’.

### 3. *Lower-Bounded Representations*

For the alternative frames in the ADP to have the same meanings, the number expressions would have to denote *complementary* quantities. Consider again, for example, frames (5) and (7) reproduced below:

<sup>6</sup> The equivalence assumption is thought to gain further empirical support from the fact that people typically *do* take the alternative frames to be equivalent when both are made available (Frisch 1993; Kühberger 1995; Mandel 2001; Stanovich and West 1998).

<sup>7</sup> See, for example, Berkeley and Humphreys 1982; Bohm and Lind 1992; Okder 2012. Also, for discussion of a related ‘attribute framing’ paradigm, see Sher and McKenzie 2006.

(5) If Program A is adopted, 200 people will be saved.

(7) If Program C is adopted, 400 people will die.

These frames can only describe the same prospect if the quantities of people being saved and dying sum to 600 (the total number of people the disease is expected to kill).<sup>8</sup> The complementarity of the number expressions in turn requires that they have ‘bilateral’ meanings, whereby ‘200’ means *no more and no less than 200*, and ‘400’ means *no more and no less than 400*. In other words, the number expressions must denote *exact* quantities.<sup>9</sup>

It has been suggested, however, that the number expressions in the ADP are often interpreted *unilaterally* —specifically, *lower-boundedly*, as denoting *minimum* quantities. For example, ‘200’ in (5) may be represented as denoting *at least* 200, and ‘400’ in (7) as denoting *at least* 400. Crucially, this ‘lower-bounding hypothesis’ stands to explain the classic risky choice framing effect. To see why, consider what happens if (5) and (7) are interpreted as in (5a) and (7a):

(5a) If Program A is adopted, at least 200 people will be saved.

(7a) If Program C is adopted, at least 400 people will die.

It is consistent with (5a) that *more than* 200 people will be saved under Program A. In contrast, (7a) ensures that *no more than* 200 people will be saved under Program C, since at least 400 of the 600 will die. Therefore, Program A has the potential to save more lives than Program C —it is a better prospect. In principle, then, this could explain the observed shift in responses, towards the sure option in the ‘be saved’ framing condition, and away from the sure option in the ‘die’ condition. Importantly, that would be a *different* explanation than the Prospect Theoretic one which, as we saw, depended on the positive wording (‘be saved’) in the first condition giving the impression of gaining, while the negative wording (‘die’) in the second condition gives the impression of losing.

<sup>8</sup> Clearly, this is only a *necessary* condition for equivalence, and is not *sufficient*. However, as mentioned at the end of the previous section, I will simply set aside other factors for now.

<sup>9</sup> Or, perhaps, *approximate* quantities —these would still be bilateral interpretations, just less precise ones. For ease of exposition, I will equate ‘bilateral’ with ‘exact’, although nothing substantive in the argument hangs on this point.

The lower-bounding hypothesis has been mooted at various times in the framing literature (going back, at least, to Macdonald 1986). However, it is only recently that it has been the subject of direct empirical investigation.<sup>10</sup>

### 3.1. Empirical Evidence

In a striking study, Mandel (2014) presents evidence that ADP-style risky-choice framing effects *can* depend on number expressions being represented lower-boundedly. In three experiments, he finds that framing effects are eliminated entirely when participants are encouraged to form bilateral ('exactly') interpretations of number expressions. Conversely, the effects are facilitated and amplified by encouraging participants to form lower-bounded ('at least') interpretations. Moreover, when participants are asked how they interpreted the number expressions, framing effects are only evident among those who report interpreting the number expressions as lower-bounded (not bilateral).

Going through the findings in a little more detail, Mandel's first experiment deployed the ADP in a within-subjects design, such that participants received *both* the 'be saved' and 'die' framing conditions.<sup>11</sup> For one group of participants, the sure options in each framing condition included the modifier 'exactly' before the number expression, as in (5b) and (7b) below.

(5b) If Program A is adopted, exactly 200 people will be saved.

(7b) If Program C is adopted, exactly 400 people will die.

For the remaining participants, the sure options contained the modifier 'at least' before the number expression (as in (5a) and (7a) above).

In a related second experiment, Mandel used a *between*-subjects design, so that each participant received the ADP in only one framing condition (the 'be saved' or 'die' condition).<sup>12</sup> In addition to the 'at least' and 'exactly' conditions, this experiment included 'no modifier'

<sup>10</sup> Some closely related studies are reported by Halberg and Teigen 2009; Teigen 2011; Teigen and Nikolaisen 2009.

<sup>11</sup> The order in which the two framing conditions were presented was counter-balanced between subjects, and all participants also received a series of unrelated, intervening 'filler' tasks.

<sup>12</sup> In both Mandel's second and third experiments, the following, modified, risky-choice scenario was used:

In a war-torn region, the lives of 600 stranded people are at stake. Two

conditions, in which the options were presented without the inclusion of ‘exactly’ or ‘at least’ before the number expression (similarly to the classic conditions of the ADP). The experiment thus included six experimental conditions in total, represented by the rows in the table below:

<i>Modifier</i>	<i>Framing condition</i>
‘Exactly’	‘Be saved’
‘Exactly’	‘Die’
‘At least’	‘Be saved’
‘At least’	‘Die’
No modifier	‘Be saved’
No modifier	‘Die’

In line with the lower-bounding hypothesis, no framing effects were observed in the ‘exactly’ conditions of either of Mandel’s first two experiments. In other words, once the modifier ‘exactly’ was added, no statistically significant difference was observed between the choices made in the ‘be saved’ and ‘die’ conditions. In contrast, the ‘at least’ conditions led to large framing effects in the standard direction: significantly larger proportions of participants chose the sure option in the ‘be saved’ condition than in the ‘die’ condition (and, conversely, significantly larger proportions chose the risky option in the ‘die’ condition than in the ‘be saved’ condition). Meanwhile, the no-modifier conditions in the second experiment yielded a medium-sized framing effect. This is consistent with the hypothesis that *some* participants were spontaneously adopting lower-bounded readings of the number

response plans with the following outcomes have been proposed. Assume that the estimates provided are accurate.

This scenario, taken from Mandel 2001, was intended to make it more plausible that the number of lives under threat could be accurately forecast. The options were then phrased similarly to those in the ADP, except that: (i) the sure options included the phrase ‘it is certain that’ before the number expressions; (ii) the risky options removed reference to ‘people’, instead referring either to ‘all 600’ or ‘nobody’; (iii) in the risky option of the ‘die’ framing condition, the order of the clauses was switched; and (iv) each of the options referred to ‘Plans’ rather than ‘Programs’. Additionally, Mandel used the following wording in the question to participants: ‘Which of the two plans would you choose —A or B?’ It seems unlikely that these relatively minor changes are responsible for the pattern of results Mandel obtains, most notably the elimination of framing effects in the ‘exactly’ conditions (particularly since framing effects *were* observed in the other conditions that used the same scenario).

expressions in these conditions (and that this is what is happening in the classic version of the ADP).<sup>13</sup>

In his third experiment, Mandel manipulated what he calls the ‘explication’ of the sure and risky options. I will focus purely on his manipulations of the *sure* options here. Mandel notes that, in the classic ADP, the sure options are only ‘partially explicated’, leaving implicit the fate of the remaining people (i.e. the 400 people who will not be saved in the ‘be saved’ condition, and the 200 who will not die in the ‘die’ condition).<sup>14</sup> Mandel formulates ‘fully explicated’ versions of the sure options as follows:

- (5c) If Plan A is adopted, it is certain that 200 people will be saved and 400 people will not be saved.
- (7c) If Plan C is adopted, it is certain that 400 people will die and 200 people will not die.

Using a between-subjects design, participants were thus divided between the four conditions represented in the rows of the table below:

<i>Framing condition</i>	<i>Explication of sure option</i>
‘Be saved’	Partial
‘Be saved’	Full
‘Die’	Partial
‘Die’	Full

Following the usual choice task, Mandel asked participants in the ‘be saved’ conditions the following question:

Did you interpret Plan A to mean (a) at most, (b) exactly, or (c) at least 200 will be saved?

In the ‘die’ condition the question read as follows:

Did you interpret Plan C to mean (a) at most, (b) exactly, or (c) at least 400 will die?

<sup>13</sup> Lending further support to this hypothesis, in the General Discussion, Mandel discusses another preliminary experiment, in which over half of participants report forming lower-bounded interpretations of the number expressions in the ADP.

<sup>14</sup> Kühberger (1995) similarly argued that the partial explication of the sure options could be responsible for risky-choice framing effects. The explication point has subsequently been developed in various ways, by Mandel 2001, Kühberger and Tanner 2010, and Tombu and Mandel 2015.

The results of this experiment showed, first, an effect of ‘explication’ on choice. Specifically, a standard framing effect arose from the classic ADP conditions, in which the sure options were only partially explicated. However, the effect was eliminated when the options were fully explicated. This suggests that the framing effect may depend on the *partial* explication of the sure options, which enables a lower-bounded interpretation.

Indeed, Mandel found an effect of explication on the interpretation of the number expressions: partially explicated sure options were correlated with participants reporting lower-bounded (‘at least’) readings of the number expressions. Conversely, fully explicated sure options were correlated with participants reporting bilateral, ‘exactly’ readings of the number expressions. This result shows how lower-bounded interpretations are linked to the sure options being only partly explicated.

Finally, there was an interaction between choice patterns and interpretations of number expressions: an overall framing effect was observed for those sets of participants reporting lower-bounded readings of the number expressions. Conversely, no framing effect was observed for the sets of participants reporting bilateral readings. This points to a direct link between lower-bounded interpretations and framing effects.

Taken together, Mandel’s results suggest that ADP-style risky-choice framing effects may be driven by some participants interpreting the number expressions in the sure options as denoting lower-bounded quantities. Mandel concludes:

In short, the findings indicate that, for most people, Options A and C in the ADP are descriptions of different events, and not merely re-descriptions of the same event. Although effects of the usual ADP positive-negative manipulation are description effects, strictly speaking, they are not framing effects, and thus should not be labeled as such. (2014, p. 1193)

It is important to note at this point that a replication attempt by Simmons and Nelson (2013) failed to confirm the results of Mandel’s second experiment. Simmons and Nelson found statistically significant framing effects, even under the ‘exactly’ conditions.<sup>15</sup> Likewise, Chick, Reyna, and Corbin (2016) still observe a framing effect when

<sup>15</sup> See Mandel (2020) for a response to Simmons and Nelson, which points out that there is no statistically significant difference between their two sets of results, despite the disparity in the statistical significance of the framing effect. Mandel also

they use the ‘exactly’ stimuli from Mandel’s second experiment.<sup>16</sup> It would be premature, then, to conclude that framing effects—even ADP-style risky-choice framing effects—depend *entirely* on lower-bounded representations of number expressions.<sup>17</sup> Nevertheless, the lower-bounding hypothesis may well provide an important, if *partial*, explanation of framing effects. In the remainder of the paper, I will focus on the issue of whether this explanation is a *rationalising* one.

### 3.2. Lower-Bounding and Rationality

Mandel argues that most of the choices made by his participants remain entirely in accord with the rational choice principle of description invariance: preferences did not shift once number expressions were made explicitly complementary via the addition of ‘exactly’; yet there was, justifiably, a shift where the number expressions were made explicitly non-complementary, via the use of ‘at least’. In Mandel’s own words: “The present findings [...] contribute to the rationality debate [...] by showing that a significant majority of participants made rational decisions by classical rational-choice criteria in traditional risky-choice framing problems” (2014, p. 1194).

While this is correct, it is important to note that the lower-bounding hypothesis does not yet provide a fully rationalising explanation of *classic* risky-choice framing effects. Instead, where number expressions are *not* modified with ‘at least’, it remains an open question whether it is rational to interpret them as denoting lower-bounded, rather than bilateral, quantities. Recall that framing researchers have standardly assumed that ‘200’ and ‘400’ in the ADP have *bilateral* meanings (such that the sure options describe the *same* outcome of 200 people being saved and 400 dying). If this is correct, people’s ‘editing’ behaviour, to arrive at *lower-bounded* interpretations of the number expressions, would still be counter-normative. They would still be making a mistake in the way they represent the options (the mistake is simply different from that which Prospect Theory attributes to them—rather than erroneously representing

notes a puzzling feature of Simmons and Nelson’s data: it suggests that participants are largely insensitive to the difference between exact and lower-bounded quantities.

<sup>16</sup> Although they use Mandel’s scenario and choice options, it is worth noting that Chick et al. introduce a new set of pre-task instructions and questionnaires, which are designed to secure bilateral readings of the number expressions even more robustly.

<sup>17</sup> Indeed, Tombu and Mandel (2015) themselves explore a separate explanation of framing effects, which appeals essentially to the (positive, negative, or mixed) valence of frames.

prospects as gains or losses, they would be erroneously representing quantities as being lower-bounded rather than bilateral).

Of course, having arrived at lower-bounded interpretations, people's subsequent 'evaluation' behaviour is now justifiable: they behave reasonably in rating the sure option more highly in the 'be saved' frame than in the 'die' frame.<sup>18</sup> Importantly, though, as discussed above, it is people's apparently irrational behaviour in the editing phase which is distinctive of framing effects (as opposed to mere *reflection* effects). Therefore, a rationalising explanation of framing effects must ultimately show that people's *representations* of prospects can be justified (and not just their subsequent evaluations). While Mandel's results provide some reason to think the number expressions in the ADP sometimes *are* interpreted lower-boundedly, they do not yet demonstrate that they *should be*. As it stands, then, the lower-bounding account falls short of rationalising framing effects.

Recall, too, that Tversky and Kahneman themselves readily acknowledge that people *do* represent outcomes differently under alternative frames. Their claim is that they *shouldn't*, if behaving perfectly rationally. Given the Prospect Theoretic account of the psychological mechanism underlying risky-choice framing effects, the burden these theorists face is to show that alternative frames obviously share a common meaning, despite the use of positive or negative linguistic expressions.<sup>19</sup> If framing effects turn out, instead, to be explained by the lower-bounding hypothesis, then the normative question turns on whether alternative frames still share a common meaning, despite the use of one or other number expression.

Mandel does indicate that claims of counter-normativity may at least be more difficult to maintain if the lower-bounding hypothesis is accepted. He appeals to linguistic evidence that number expressions can give rise to lower-bounded interpretations in certain contexts (2014, p. 1186). However, Mandel explicitly defers further discussion of *which* contexts licence lower-bounded interpretations. He writes:

<sup>18</sup> Perhaps, then, there are two notions of rationality at play, linked respectively to people's editing behaviour (i.e. their interpretation of linguistic stimuli) and their evaluation behaviour (i.e. their choice between the two options). I am grateful to David Mandel for suggesting this and, although I cannot pursue the point here, I believe it would be another interesting area of research for empirically informed philosophy.

<sup>19</sup> Again, this is debatable on various grounds but I will not pursue each one here.



Although a specification of how various types of context (e.g., conversational, semantic, and sentential) affect the interpretation of [numeric quantifiers] is beyond the aims of this article, the present research [...] suggests the need for a comprehensive psycholinguistic account of numeric quantifier use and interpretation. (2014, p. 1196)

To reinforce Mandel's point here, if the lower-bounding hypothesis is to support a genuinely rationalising explanation of risky-choice framing effects, it will first be necessary to resolve the question of how the number expressions in those paradigms *should* be interpreted—in other words, what they *mean*. Mandel may recognise that this is needed. However, others have interpreted him as straightforwardly offering a rationalising explanation—including, for example, Chick et al. 2016; McKenzie, Sher, Leong, and Müller-Trede 2018. Given that the question of rationality is—and has always been—at the heart of debates about framing, it is important to understand precisely which linguistic foundations could ground a genuinely rationalising incarnation of the lower-bounding account.

#### 4. *The Meanings of Number Expressions*

If the lower-bounding hypothesis is to provide a genuinely rationalising explanation of risky-choice framing effects, it must be shown that experimental participants are rational to represent the number expressions as denoting lower-bounded quantities. Specifically, it must be shown that such interpretations are licensed, on the basis of the expressions having genuinely lower-bounded *meanings* in the relevant contexts. In this section, I briefly survey various possible arguments to that effect, drawing on an ongoing debate about the semantics and pragmatics of number expressions.<sup>20</sup> Although I will not seek to provide a resolution, I will indicate how each position could potentially support a rationalising account of framing effects. Note that none of them actually does so as yet. Jumping ahead for a moment, in section 5 I will argue that choosing *between* the positions, in pursuit of the requisite justification, will ultimately require deeper philosophical engagement.

While it is widely recognised that number expressions obtain lower-bounded readings in certain contexts, some theorists have sought to trace this back to the expressions' conventional semantics, while others have attributed it to their pragmatic enrichment in particular contexts of use. I begin by discussing the traditional

<sup>20</sup> For a fuller survey of this debate, see Spector 2013.

neo-Gricean view, which has been put forward, for example, by Horn (1989) and Levinson (2000). They take number expressions to be lower-bounded as a matter of conventional meaning. On the neo-Gricean view, then, ‘200’ conventionally denotes *at least 200*, while ‘400’ conventionally denotes *at least 400*. If correct, this would render sentences like (5) and (7) inequivalent to one another at a basic semantic level. While the theorists in this camp readily acknowledge that number expressions are sometimes *interpreted* as denoting bilateral (or *upper*-bounded) quantities, this is thought to be the result of pragmatic modulation of the conventional meaning, in light of wider contextual considerations.

If number expressions have conventionally lower-bounded meanings, that might already be thought sufficient to license their being represented as lower-bounded. For example, according to some semantic theories, conventional meanings are *always* psychologically realised during successful linguistic communication.<sup>21</sup> Of course, showing that the conventional meanings should also be deployed in making evaluative judgements would be a further step.<sup>22</sup> Indeed, participants’ choice behaviour might still be thought to violate rational choice principles. This is because it is unclear whether they ought to evaluate the options on the basis of (conventionally-derived) lower-bounded representations, or (pragmatically-derived) bilateral ones. Consider, for example, how pragmatic enrichment of the number expression might be expected to proceed on a standard Gricean analysis:<sup>23</sup> given a statement of (5), if *more than 200* people would be saved under Program A, then presumably that would have been stated explicitly—at least if we assume that the number of people who will be saved is *known*, and that the information being provided is *accurate*, *relevant*, and so on. Since (5) does *not* state that more than 200 people will be saved, it may be legitimate to infer that

<sup>21</sup> See, for example, Cappelen and Lepore 2005.

<sup>22</sup> In general, there will always be a further step to argue that a representation arrived at during the ‘editing’ phase should be deployed in the ‘evaluation’ phase. However, it is first necessary to demonstrate that a lower-bounded interpretation can reasonably and justifiably be *derived*; if not, there would be no hope of giving the lower-bounding hypothesis a rational basis. I am grateful to an anonymous reviewer for pressing me on this point.

<sup>23</sup> I intend to remain neutral here on whether audiences actually reason in this way. However, this is a question which has received considerable attention in the literature on Gricean pragmatics (see, for example, Geurts and Rubio-Fernández 2015; Saul 2002).

*no more than* 200 will be saved—in other words, that *exactly* 200 people will be saved.<sup>24</sup>

Given this Gricean rationale, is it still possible to justify a lower-bounded interpretation of the number expression (and an evaluation based on that interpretation), rather than a bilateral one? Perhaps. After all, it is not obvious that the number of people who will be saved is really *known*; the vignette explicitly states that the figures are *estimates*. Therefore, the reason for failing to state that more than 200 people will be saved could be a purely epistemic one—it is simply unknown how many people beyond the first 200 will eventually be saved. On that basis, participants might be rational to stick with the conventional lower-bounded meaning of the number expression, leaving open the possibility that more than 200 people will be saved, rather than adopting a pragmatically-enriched, bilateral interpretation.<sup>25</sup>

On the other hand, participants are explicitly instructed to assume that the options provide *exact* scientific estimates of the consequences. That might be considered to give them a good reason to adopt bilateral interpretations of the number expressions. Ultimately, then, whether or not the deployment of lower-bounded interpretations can be justified—and the framing effects that result—will depend on precisely how the discourse context functions to constrain meaning. As Mandel indicates in the earlier quote, this is an issue for wider research in linguistic pragmatics to address. For now, then, a rationalising explanation remains possible but not secure.

A distinct approach to the conventional semantics of number expressions takes these expressions to be *ambiguous*, or *polysemous*, having multiple lexical entries, or senses. Geurts (2006) develops an analysis along these lines. He argues that number expressions have both bilateral and lower-bounded senses (although he takes the bilateral senses to be more basic and generally dominant).

Adopting an ambiguity or polysemy view of number expressions, one would then need to show that it is *reasonable* to adopt lower-bounded interpretations of the number expressions in the ADP. As before, the task of showing that is made somewhat harder by

<sup>24</sup> And, of course, a similar explanation could be provided for the alternative frame in (7).

<sup>25</sup> Moreover, it is perhaps notable that (5) fails to state explicitly how many people will *not* be saved, whereas the risky option in (6) does so. This point of contrast between the two options might indicate that there is some uncertainty about the outcome of Program A, further justifying a lower-bounded interpretation of the number expression.

the fact that participants are told that the options provide *exact* scientific estimates, which might be expected to prime a bilateral interpretation (and, if that interpretation is also the dominant one, as Geurts suggests, then selecting the lower-bounded sense would seem especially odd). Only if a satisfactory explanation can be given here would risky-choice framing effects be consistent with the principle of description invariance. In the end, this will depend on our theory of how alternative senses should be selected in context.

Another approach claims that number expressions have *underspecified* conventional semantics. In other words, the expressions themselves have neither bilateral nor lower-bounded denotations. This view has been advocated by Kempson and Cormack (1981) and Carston (1988, 1998). They argue that, on each occasion of use, the underspecified conventional meaning of a number expression must be pragmatically enriched in one way or other, so that the audience arrives at a lower-bounded or bilateral (or, potentially, *upper-bounded*) interpretation.<sup>26</sup>

<sup>26</sup> In this regard, it is interesting to consider the additional study described by Mandel in the General Discussion of his 2014 paper. This study aims to shed light on some empirical results obtained by Jou, Shanteau, and Harris (1996). Jou et al. find that risky-choice framing effects can be eliminated by providing a rationale for the outcomes and probabilities associated with each option. For example, their version of the ADP states that:

Without treatment, a person who has contracted the disease will die in a few days. Six hundred people have been diagnosed as having contracted the disease. Some substance, extracted from living human organs and extremely difficult to obtain, can cure the disease. Unfortunately, there is only enough of this substance for 200 people.

Mandel hypothesises that this encourages recipients to interpret the sure option in the ‘be saved’ condition as saving *at most* 200 people (since the quantity of the curative substance is known to be upper-bounded), whereas the sure option in the ‘die’ condition is still most naturally understood to result in *at least* 400 people dying. (Arguably, though, in Jou et al.’s original experiment, ‘exact’ interpretations are equally justified, since the task instructions state: “Your estimate of the likelihood of the occurrence of some event should be strictly based on the probabilities given you in each problem, not based on your personal intuitions or experience about the likelihood of a certain event’s occurring”; however, I will ignore this wrinkle for now). Taken together, the two interpretations —‘at most 200’ and ‘at least 400’— would make the sure option in the ‘be saved’ frame just as unappealing as it is in the ‘die’ frame, thus explaining the absence of a framing effect. In line with the hypothesis, participants in Mandel’s initial study generally did reject an ‘exactly’ interpretation of ‘200’, subsequently endorsing an ‘at most’ interpretation. And they tended to reject an ‘exactly’ interpretation of ‘400’, subsequently endorsing an ‘at least’ interpretation. This is further evidence that different interpretations of the number expressions can be contextually induced. It would be too quick, though, to

Given an underspecification view, the lower-bounding hypothesis would provide a rationalising explanation of risky-choice framing effects just if participants could be justified in enriching number expressions lower-boundedly in those contexts. Whether that is so will depend on precisely how such enrichment is supposed to proceed, i.e. which contextual factors may be taken into account, and how they are to be integrated within the interpretation process. That, in turn, will require a rich, psychological theory of interpretation.<sup>27</sup>

The final account I canvas here has been advocated by Breheny (2008), who analyses number expressions as having straightforwardly *bilateral* conventional meanings. Breheny recognises that such expressions may obtain lower-bounded interpretations in certain contexts, but this is taken to be the result of pragmatic enrichment (specifically, he appeals to the role of background knowledge in generating unilateral meanings). In this sense, Breheny's position is diametrically opposed to the neo-Gricean approach discussed above.

On a Breheny-style account, the question of whether lower-bounded interpretations are *rational* would depend, again, on our final pragmatic theory (including how individuals' background knowledge should be brought to bear during interpretation). So long as participants can be justified in deriving enriched, lower-bounded interpretations of the number expressions in the ADP, they would not be violating the principle of description invariance; there would be an important sense in which the number expressions actually *have* lower-bounded meanings in that context.

It is clear from the preceding discussion that the ability of the approaches described here to support a fully rationalising explanation of framing effects will depend on our final theory of one or more aspects of linguistic communication. In the meantime, the jury is

draw any conclusions about the semantics of number expressions. In particular, the fact that participants tend to interpret '200' as meaning *at most 200* in the context of this scenario does not show that number expressions are semantically underspecified. It remains possible that people can still distinguish between the literal semantics of the expression (whether lower-bounded, exact, or ambiguous between the two) and information inferable from the wider context (i.e. that there is only enough of the substance for 200 people and that, although it 'can' cure the disease, perhaps it will not turn out to be successful for all 200). The suggestion, then, is that Mandel's experiment successfully probed pragmatic interpretations, without necessarily telling us anything about the semantics of number expressions themselves.

<sup>27</sup> According to the Relevance Theoretic approach endorsed by Carston, the lower-bounded interpretation would need to have a sufficiently important effect on a participant's beliefs to justify his/ her cognitive effort in deriving it (Carston 2002; Sperber and Wilson 1995).

out on whether participants ultimately behave rationally in deriving and deploying lower-bounded representations of number expressions during risky-choice framing tasks. At most, each approach explains *how* participants arrive at those interpretations (semantically or pragmatically), but not whether they are the best interpretations to adopt, all things considered. That said, it is equally clear that it would be premature to conclude that framing effects are *counter*-normative. Thus, the lower-bounding hypothesis at least leaves in play the possibility that framing effects are rational. The importance of this result should not be downplayed, given how widespread and influential the opposite conclusion has become.

The contribution of the discussion in this section has been to outline various ways in which Mandel's empirical analysis could potentially be deepened, by appeal to a foundational account of the meanings of number expressions. One such account will ultimately be required if the lower-bounding hypothesis is to support a fully rationalising explanation of framing effects. Having established this, the point I want to emphasise in the next section of the paper is that deciding *between* these foundational positions will require empirically-engaged philosophy.

### 5. *The Semantics-Pragmatics Divide*

The question of what number expressions mean, at semantic and pragmatic levels, depends in large part on how we define the concepts of *semantics* and *pragmatics*, and where we draw the boundary between the two. This is a project which has exercised philosophers of language for many years. Although we are, of course, free to stipulate how technical terminology is used, presumably an important objective is for the concepts of semantics and pragmatics to have genuine explanatory value, and this should guide their definition and distinction. Even granting this point, though, the semantics-pragmatics boundary is unlikely to fall out automatically from pure empirical inquiry, since we must first agree *which* real-world phenomena we are trying to explain. For illustration, below I briefly consider three metaphysical distinctions we may want the semantics-pragmatics divide to track (and there are almost certainly others). Depending on how this conceptual issue is resolved, it may prove possible to endorse one of the approaches described in section 4, thus providing the basis for a genuinely rationalising explanation of framing effects.

First, the purpose of semantics could be thought of as tracking certain *conventional* effects on interpretation, which depend on arbitrary

features of language, such as the fact that the word ‘fire’ in English means *fire*. Meanwhile, pragmatics could track meanings that emerge from non-conventional contextual effects: for example, the sentence ‘There’s a fire’ is used to vastly different effect when it is uttered to other passengers in a crowded underground station, as opposed to one’s friends as they walk into a warm pub. Lepore and Stone (2015) have recently elaborated a view which treats conventionality as the hallmark of semantics (but, in the process, absorbs all kinds of interpretive effects into the category of semantics, which have not traditionally been considered part of conventional meaning).

Others have proposed a *psychological* grounding for semantics and pragmatics, seeking to map them to distinct kinds of cognitive process. According to Relevance Theory, for example, semantic processing involves the decoding of linguistic stimuli, whereas pragmatic processing is governed by an entirely different cognitive efficiency principle (Carston 2002; Sperber and Wilson 1995). Borg also maps semantics and pragmatics to different (modular vs. non-modular) psychological systems (Borg 2004, 2012) —although it should be noted that her view of the contextual effects on meaning diverges importantly from that of Relevance Theorists.

Another way to draw the boundary would be by appeal to normative criteria derived from our everyday communicative practices. The idea here would be that semantics could track those meanings that are hardest for speakers to retract, and to which they remain strictly committed. In contrast, pragmatic meanings may be relatively easy to cancel or deny. Saul (2013) adopts this kind of approach in mapping the difference between lying and misleading to the difference between what is ‘said’ and what is conveyed less directly. See also Borg 2017 for discussion of how speakers’ liabilities track semantic and pragmatic differences.

If we are lucky, each of these ways of understanding semantic and pragmatic meaning will be perfectly aligned, sorting communicated contents in exactly the same way into one basket or the other. Perhaps more likely, though, we will need to make explicit decisions about which distinction to track. Such decisions are ones to which philosophers of language are particularly well-equipped to contribute.

The question of how to draw the semantics-pragmatics boundary will feed directly into the analysis of the meaning of number expressions. For example, if semantics and pragmatics are held to track distinct cognitive processes, we would expect the semantic and pragmatic meanings of number expressions to have different psychological signatures (perhaps being derived at different speeds, involving

different regions of the brain, or being susceptible to different kinds of interference). Alternatively, if we adopt a normative criterion, the semantics and pragmatics of number expressions ought to reflect our ordinary communicative practices. Therefore, in order to evaluate the views surveyed in the previous section, we need to establish *which* real-world phenomena the posited meanings should be tracking.

The upshot of this (necessarily cursory) discussion, then, is that distinguishing and deciding between the defining criteria of semantic and pragmatic meanings is essential for choosing between analyses of the meanings of number expressions. As we saw, settling on one of these analyses will, in turn, be a necessary step towards rendering the lower-bounding hypothesis a *rationalising* explanation of risky-choice framing effects. In this way, then, further exploration of an empirical debate in psychological framing research has led us inexorably to a foundational philosophical question concerning the nature of semantic and pragmatic meaning. Looking back in the other direction, it is clear that empirically informed discussions in the philosophy of language have demonstrable application in advancing framing research and its associated real-world impacts.

## 6. Conclusion

In the case of risky-choice framing effects —and, in particular, the hypothesis that lower-bounded interpretations of number expressions can explain these effects— empirical and philosophical inquiries are importantly intertwined. I have argued that competing explanations of the representation and evaluation of alternative frames depend, ultimately, on foundational issues in the philosophy of language, concerning semantics, pragmatics, and the interface between the two. In doing so, I hope to have indicated how, by reaching across disciplinary boundaries, philosophers can shed light on the conceptual questions pervading empirical work, and can contribute positively to the development of scientific research. Of course, it remains to be seen how representative this case study is of wider epistemic inquiry. However, it would perhaps not be too surprising to discover that science and philosophy are profoundly and pervasively entangled across the piece.<sup>28</sup>

<sup>28</sup> This work was supported by the Arts and Humanities Research Council, via the South, West and Wales Doctoral Training Partnership (grant number AH/L503939/1). I would like to thank Emma Borg, Nat Hansen, and David Mandel for invaluable feedback on earlier versions of the paper. I am also extremely grateful for comments provided by *Crítica*'s editorial board and reviewers.



## REFERENCES

- Berkeley, D. and P. Humphreys, 1982, “Structuring Decision Problems and the ‘Bias Heuristic’”, *Acta Psychologica*, vol. 50, no. 3, pp. 201–252. (doi: 10.1016/0001-6918(82)90042-7)
- Bermúdez, J.L., 2018, “Frames, Rationality, and Self-Control”, in J.L. Bermúdez (comp.), *Self-Control, Decision Theory, and Rationality*, Cambridge University Press, Cambridge.
- , 2009, “The Second Challenge: Individuating Outcomes”, in J.L. Bermúdez (comp.), *Decision Theory and Rationality*, Oxford University Press, Oxford.
- Bohm, P. and H. Lind, 1992, “A Note on the Robustness of a Classical Framing Result”, *Journal of Economic Psychology*, vol. 13, no. 2, pp. 355–361. (doi:10.1016/0167-4870(92)90039-A)
- Borg, E., 2017, “Explanatory Roles for Minimal Content”, *Noûs*, vol. 53, no. 3. (doi:10.1111/nous.12217)
- , 2012, *Pursuing Meaning*, Oxford University Press, Oxford.
- , 2004, *Minimal Semantics*, Clarendon Press, Oxford.
- Breheny, R., 2008, “A New Look at the Semantics and Pragmatics of Numerically Quantified Noun Phrases”, *Journal of Semantics*, vol. 25, no. 2, pp. 93–139. (doi: 10.1093/jos/ffm016)
- Cappelen, H. and E. Lepore, 2005, *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*, Wiley-Blackwell, Oxford.
- Carston, R., 2002, *Thoughts and Utterances: The Pragmatics of Explicit Communication*, Blackwell, Oxford.
- , 1998, “Informativeness, Relevance and Scalar Implicature”, in R. Carston and S. Uchida (comps.), *Relevance Theory: Applications and Implications*, John Benjamins Publishing Company, Amsterdam.
- , 1988, “Implicature, Explicature, and Truth-Theoretic Semantics”, in R.M. Kempson (comp.), *Mental Representations: The Interface between Language and Reality*, Cambridge University Press, Cambridge.
- Chick, C.F., Reyna, V.F. and J.C. Corbin, 2016, “Framing Effects Are Robust to Linguistic Disambiguation: A Critical Test of Contemporary Theory”, *Journal of Experimental Psychology: Learning Memory and Cognition*, vol. 42, no. 2, pp. 238–256. (doi:10.1037/xlm0000158)
- Frisch, D., 1993, “Reasons for Framing Effects”, *Organizational Behavior and Human Decision Processes*, vol. 54, no. 3, pp. 399–429. (doi:10.1006/obhd.1993.1017)
- Geurts, B., 2006, “Take ‘Five’: the Meaning and Use of a Number Word”, in S. Vogeeler and L. Tasmowski (comps.), *Non-Definiteness and Plurality*, John Benjamins Publishing Company, Amsterdam.
- Geurts, B. and P. Rubio-Fernández, 2015, “Pragmatics and Processing”, *Ratio*, vol. 28, no. 4, pp. 446–469. (doi: 10.1111/rati.12113)

- Grice, H.P., 1989, *Studies in the Way of Words*, Harvard University Press, Cambridge, MA.
- Halberg, A.-M. and K.H. Teigen, 2009, "Framing of Imprecise Quantities: When Are Lower Interval Bounds Preferred to Upper Bounds?", *Journal of Behavioral Decision Making*, vol. 22, no. 5, pp. 490–509. (doi:10.1002/bdm.635)
- Horn, L., 1989, *A Natural History of Negation*, University of Chicago Press, Chicago.
- Jou, J., J. Shanteau and R.J. Harris, 1996, "An Information Processing View of Framing Effects: The Role of Causal Schemas in Decision Making", *Memory and Cognition*, vol. 24, no. 1, pp. 1–15. (doi:10.3758/BF03197268)
- Kahneman, D. and A. Tversky, 1979, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, vol. 47, no. 2, pp. 263–291.
- Kempson, R.M. and A. Cormack, 1981, "Ambiguity and Quantification", *Linguistics and Philosophy*, vol. 4, no. 2, pp. 259–309. (doi: 10.1007/BF00350141)
- Kühberger, A., 1998, "The Influence of Framing on Risky Decisions: A Meta-analysis", *Organizational Behavior and Human Decision Processes*, vol. 75, no. 1, pp. 23–55. (doi:10.1006/obhd.1998.2781)
- , 1995, "The Framing of Decisions: A New Look at Old Problems", *Organizational Behavior and Human Decision Processes*, vol. 62, no. 2, pp. 230–240. (doi: 10.1006/obhd.1995.1046)
- Kühberger, A., M. Schulte-Mecklenbeck and J. Perner, 1999, "The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks", *Organizational Behavior and Human Decision Processes*, vol. 78, no. 3, pp. 204–231. (doi:10.1006/obhd.1999.2830)
- Kühberger, A. and C. Tanner, 2010, "Risky Choice Framing: Task Versions and a Comparison of Prospect Theory and Fuzzy-Trace Theory", *Journal of Behavioral Decision Making*, vol. 23, no. 3, pp. 314–329. (doi:10.1002/bdm.656)
- Lepore, E. and M. Stone, 2015, *Imagination and Convention: Distinguishing Grammar and Inference in Language*, Oxford University Press, Oxford.
- Levin, I.P., S.L. Schneider and G.J. Gaeth, 1998, "Not All Frames Are Created Equal: A Typology and Critical Analysis of Framing Effects", *Organizational Behavior and Human Decision Processes*, vol. 76, no. 2, pp. 149–188. (doi: 10.1006/obhd.1998.2804)
- Levinson, S.C., 2000, *Presumptive Meanings: The Theory of Generalized Conversational Implicature*, MIT Press, Cambridge, MA.
- Macdonald, R.R., 1986, "Credible Conceptions and Implausible Probabilities", *The British Journal of Mathematical and Statistical Psychology*, vol. 39, no. 1, pp. 15–27. (doi:10.1111/j.2044-8317.1986.tb00842.x)
- Mandel, D.R., 2020, "Framing Effects, Replications, and Scientific Inference: Reply to Commentary by Simmons and Nelson (2013) on Mandel

- (2014)", <https://doi.org/10.31234/osf.io/34jeg> [date of last consultation: 30/08/2020].
- Mandel, D.R., 2014, "Do Framing Effects Reveal Irrational Choice?", *Journal of Experimental Psychology. General*, vol. 143, no. 3, pp. 1185–1198. (doi:10.1037/a0034207)
- , 2001, "Gain-Loss Framing and Choice: Separating Outcome Formulations from Descriptor Formulations", *Organizational Behavior and Human Decision Processes*, vol. 85, no. 1, pp. 56–76. (doi:10.1006/obhd.2000.2932)
- McKenzie, C.R.M., S. Sher, L.M. Leong and J. Müller-Trede, 2018, "Constructed Preferences, Rationality, and Choice Architecture", *Review of Behavioral Economics*, vol. 5, no. 3–4, pp. 337–360. (doi: 10.1561/105.0000091)
- Okder, H., 2012, "The Illusion of the Framing Effect in Risky Decision Making", *Journal of Behavioral Decision Making*, vol. 25, no. 1, pp. 63–73. (doi:10.1002/bdm.715)
- Saul, J.M., 2013, *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics*, Oxford University Press, Oxford.
- , 2002, "What Is Said and Psychological Reality; Grice's Project and Relevance Theorists' Criticisms", *Linguistics and Philosophy*, vol. 25, no. 3, pp. 347–372. (doi:10.1023/A:1015221313887)
- Sher, S. and C.R.M. McKenzie, 2006, "Information Leakage from Logically Equivalent Frames", *Cognition*, vol. 101, no. 3, pp. 467–494. (doi:10.1016/j.cognition.2005.11.001)
- Simmons, J. and L. Nelson, 2013, "'Exactly': The Most Famous Framing Effect is Robust to Precise Wording", <http://datacolada.org/11> [date of last consultation: 30/08/2020].
- Spector, B., 2013, "Bare Numerals and Scalar Implicatures", *Language and Linguistics Compass*, vol. 7, no. 5, pp. 273–294. (doi:10.1111/lnc3.12018)
- Sperber, D. and D. Wilson, 1995, *Relevance: Communication and Cognition* (2nd ed.), Blackwell, Oxford.
- Stanovich, K.E. and R.F. West, 1998, "Individual Differences in Framing and Conjunction Effects", *Thinking and Reasoning*, vol. 4, no. 4, pp. 289–317. (doi:10.1080/135467898394094)
- Steiger, A. and A. Kühberger, 2018, "A Meta-Analytic Re-Appraisal of the Framing Effect", *Zeitschrift für Psychologie*, vol. 226, no. 1, pp. 45–55. (doi:10.1027/2151-2604/a000321)
- Teigen, K.H., 2011, "When Frames Meet Realities: On the Perceived Correctness of Inaccurate Estimates", in G. Keren (comp.), *Perspectives on Framing*, Psychology Press, Abingdon, Oxfordshire.
- Teigen, K.H. and M.I. Nikolaisen, 2009, "Incorrect Estimates and False Reports: How Framing Modifies Truth", *Thinking and Reasoning*, vol. 15, no. 3, pp. 268–293. (doi:10.1080/13546780903020999)

- Tombu, M. and D.R. Mandel, 2015, “When Does Framing Influence Preferences, Risk Perceptions, and Risk Attitudes? The Explicated Valence Account”, *Journal of Behavioral Decision Making*, vol. 28, no. 5, pp. 464–476. (doi:10.1002/bdm.1863)
- Tversky, A. and D. Kahneman, 1981, “The Framing of Decisions and the Psychology of Choice”, *Science*, vol. 211, no. 4481, pp. 453–458. (doi: 10.1126/science.7455683)

*Received: February 27, 2020; revised: June 8, 2020; accepted: August 28, 2020.*