



Framing Effects and Fuzzy Traces: ‘Some’ Observations

Sarah A. Fisher¹ 

Accepted: 27 April 2021 / Published online: 10 May 2021
© The Author(s) 2021, corrected publication 2021

Abstract

Framing effects occur when people respond differently to the same information, just because it is conveyed in different words. For example, in the classic ‘Disease Problem’ introduced by Amos Tversky and Daniel Kahneman, people’s choices between alternative interventions depend on whether these are described positively, in terms of the number of people who will be saved, or negatively in terms of the corresponding number who will die. In this paper, I discuss an account of framing effects based on ‘fuzzy-trace theory’. The central claim of this account is that people represent the numbers in framing problems in a ‘gist-like’ way, as ‘some’; and that this creates a categorical contrast between ‘some’ people being saved (or dying) and ‘no’ people being saved (or dying). I argue that fuzzy-trace theory’s gist-like representation, ‘some’, must have the semantics of ‘some and possibly all’, not ‘some but not all’. I show how this commits fuzzy-trace theory to a modest version of a rival ‘lower bounding hypothesis’, according to which lower-bounded interpretations of quantities contribute to framing effects by rendering the alternative descriptions extensionally inequivalent. As a result, fuzzy-trace theory is incoherent as it stands. Making sense of it requires dropping, or refining, the claim that decision-makers perceive alternatively framed options as extensionally equivalent; and the related claim that framing effects are irrational. I end by suggesting that, whereas the modest lower bounding hypothesis is well supported, there is currently less evidence for the core element of the fuzzy trace account.

1 Introduction

Consider the following hypothetical scenario, introduced by Tversky and Kahneman (1981) and subsequently used in a plethora of framing studies:

- 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

✉ Sarah A. Fisher
sarah.fisher@univie.ac.at

¹ University of Vienna, Vienna, Austria

been proposed. Assume the exact scientific estimates of the consequences are as follows:

In one condition, participants then receive the following options:

- 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, participants receive the following options:

- 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.

Participants in each condition are asked which of the options they would favor. Tversky and Kahneman find that most of the experimental participants in the ‘be saved’ condition favor Program A (72% select this option, compared with only 28% who favor Program B). Meanwhile, in the ‘die’ condition, most of the participants favor Program D (78% select this option, compared with just 22% who favor Program C). Crucially, this is despite the fact that Programs A and C are supposed to be extensionally equivalent, each resulting in 200 people being saved and 400 dying. Programs B and D are also supposed to be extensionally equivalent, both carrying a one-third probability of 600 people being saved/ none dying, and a two-thirds probability of 600 people dying/ none being saved.¹ The difference between the two experimental conditions, then, is purely in how the options are described. Nevertheless, as Tversky and Kahneman’s data show, people tend to prefer the *sure option* (Program A) in the ‘be saved’ condition; but they prefer the *risky option* (Program D) in the ‘die’ condition.

The reversal in preferences induced by the Disease Problem is known as a ‘framing effect’.² Indeed, this remains the central case in what is now a large literature on framing. Note that I will understand framing effects to encompass *shifts* in preferences, as well as wholesale *reversals*. For example, even if most participants had preferred the sure option in both framing conditions of the Disease Problem, as long as that option was preferred by a *larger* majority in the positively-worded condition, that would suffice for a framing effect to be present.³ Framing effects can thus be thought of as *systematic shifts in preferences between alternative options, due to how the options are worded*.

Various explanations of this puzzling phenomenon have been proposed. Tversky and Kahneman advance an explanation of framing effects based on their ‘prospect theory’ of

¹ These equivalences might be called into question in various ways, one of which will be discussed in detail below. For now, though, let us assume that they hold.

² Specifically, it is an example of a ‘risky choice’ framing effect, according to the typology of Levin et al. (1998). For now, I set aside the other framing paradigms discussed there.

³ A similar strategy is adopted by Levin et al. (1998) and Kühberger (1998), among others. The distinction between reversals and shifts is also discussed by Maule (1989) and Wang (1996) (who labels these ‘bidirectional’ and ‘unidirectional’ framing effects).

decision-making under risk (Kahneman and Tversky 1979). Prospect theory is intended to provide a descriptively adequate model of decision-making, which departs in various ways from the normative principles embodied in classical 'expected utility theory'. For current purposes, the important feature of prospect theory is the idea that people's risk attitudes are sensitive to whether utility is being *gained* or *lost*. In particular, people are hypothesized to be *risk-averse* in the domain of gains (preferring sure gains to risky gains that have the same expected value) but *risk-seeking* in the domain of losses (preferring risky losses to sure losses that have the same expected value).

Returning to the Disease Problem, this idea is extended to cases where the prospects merely *seem* like gains or losses, despite their actually having the same overall impact on utility (Tversky and Kahneman 1981). Where the options are framed positively, in terms of the number of people who will 'be saved', they sound like gains. Faced with apparent gains, it is hypothesized that people tend to be risk-averse, preferring the sure option to the risky option (despite both having the same expected value). Conversely, where the options are framed negatively, in terms of the number of people who will 'die', they sound like losses. Faced with apparent losses, it is hypothesized that people tend to be risk-seeking, preferring the risky option to the sure option. This approach could potentially explain the 'shifty' preferences that constitute the framing effect.

In what follows I will focus on a different explanation of framing effects, which has emerged in the subsequent literature. This account, based on 'fuzzy-trace theory', appeals to 'gist-like' representations of the numbers involved in the Disease Problem. I will make an important conceptual clarification regarding the account's central claim that the numbers are represented as 'some'. Specifically, I will show that 'some' here must have the lower-bounded semantics of 'some and possibly all', rather than the bilateral semantics of 'some but not all'. Crucially, this commits fuzzy-trace theory to a modest version of a rival 'lower bounding hypothesis', according to which lower-bounded interpretations of quantities contribute to framing effects (without explaining them entirely) by rendering the alternative descriptions extensionally inequivalent. As a result, fuzzy-trace theory is incoherent as it currently stands. Making sense of it requires dropping, or refining, the claim that decision-makers perceive alternatively framed options as extensionally equivalent; and the related claim that framing effects are irrational. I will end by suggesting that, whereas the modest lower bounding hypothesis is well supported, there is currently less evidence for the core element of the fuzzy trace account, which concerns a categorical contrast between 'some' and 'none'.

The structure of the paper is as follows: in Section 2 I introduce fuzzy-trace theory and explain how it seeks to explain framing effects. In Section 3 I clarify the semantics of fuzzy-trace theory's 'gist-like' representation, 'some', arguing that it must be 'some and possibly all', not 'some but not all'. In Section 4 I describe a rival account of framing effects, the 'lower bounding hypothesis'. In Section 5 I argue that fuzzy-trace theory is committed to a modest version of that hypothesis, and I describe how the theory must be modified, if it is to be internally consistent. Crucially, this requires dropping, or refining, the claim that decision-makers perceive alternatively framed options as extensionally equivalent; and the related claim that framing effects are irrational. In Section 6, I suggest that there is currently little evidence to suggest that framing effects depend on fuzzy traces. I conclude by proposing two avenues for future empirical work.

2 Fuzzy-Trace Theory

Fuzzy-trace theory (henceforth FTT) was originally developed in the mid-1980s in the context of memory research (Brainerd and Kingma 1984, 1985). Its central claim is that people simultaneously encode information in precise verbatim ways and in more ‘gist-like’ ways. While FTT is a broad theory, with potentially far-reaching implications across various domains of cognitive science, my focus here will be on its application to framing effects. Of particular relevance to this discussion, the proponents of FTT put forward an account of how *numbers* are represented during numerical reasoning tasks, including framing problems (Broniatowski and Reyna 2018; Chick et al. 2016; Reyna and Brainerd 1991, 2011; Reyna et al. 2014). To see how this works, let us return to the Disease Problem. The fuzzy-trace theorists propose that number expressions, like ‘200’ and ‘400’ in the sure options of that problem, activate multiple mental representations, ranging along a continuum of precision. At one extreme is the verbatim representation, i.e. ‘exactly 200’ or ‘exactly 400’. At the other extreme is the most ‘gist-like’ representation – simply ‘some’.

According to FTT, thinkers preferentially deploy the most gist-like representation compatible with completing the task at hand.⁴ So FTT predicts that, wherever possible, people will operate with representations of ‘200’ and ‘400’ (and any other number) as ‘some’.⁵ Of course, operating with the gist-like representation, ‘some’, will not always be sufficient: for example, if one is judging whether it is better to have £200 or £300 then comparing ‘some pounds’ with ‘some pounds’ won’t help. In such scenarios, thinkers must ascend the precision hierarchy, deploying incrementally more precise representations, as required. It is predicted that verbatim ‘exactly’ representations are used only as a last resort. As we will see next, the Disease Problem can be solved, at least in principle, without needing to move beyond the most-gist like representations.⁶

⁴ Reyna et al. (2014, p. 77) suggest that this preference is reflected in a sequential process:

[T]he decision-making process begins with the simplest qualitative distinctions (a fuzzy-processing preference); decision makers rely on finer distinctions only if the simplest gist representations of the options are identical.

Later developments of the theory, however, point towards a parallel processing picture (Broniatowski and Reyna 2018). For ease of exposition, I will assume that processing is sequential. However, nothing in the argument hangs on this point. The important claim is the more general one that Chick et al. put as follows:

[F]uzzy trace theory predicts that individuals base their decision on the lowest level gist representation sufficient to produce a decision (Chick et al. 2016, p. 240).

⁵ Note that FTT does not merely hypothesize that number expressions are interpreted *approximately* (such that, for example, ‘200’ is represented as ‘about 200’). Instead, it claims that the quantity is represented as ‘some’. The account of framing effects depends on this claim.

⁶ One might question the psychological plausibility of various aspects of FTT, including whether number expressions are really transformed into gist-like representations; or whether it is really easier to reason with gist-like representations than verbatim ones. I am grateful to David Mandel (p.c.) for raising these points. For reasons of space, I will not attempt to explore them further here. However, see Section 6 for some related discussion.

FTT seeks to explain the framing effect induced by the Disease Problem in terms of a categorical contrast between 'some' people and 'no' people. In the 'be saved' condition, the contrast is between 'some' people being saved under Program A and 'no' people being saved under Program B. Specifically, FTT predicts that the options are represented in the ways presented below (Reyna and Brainerd 2011, p. 188; Reyna et al. 2014, p. 77). First, the sure option in the 'be saved' condition is represented as in (1a):

- (1a) If Program A is adopted, some people will be saved.

In (1a) '200' has been substituted by 'some' in line with FTT's hypothesis that people prefer to reason with the most 'gist-like' representation.

Meanwhile, the risky option in the 'be saved' condition is thought to be represented as in (2a):

- (2a) If Program B is adopted, there is some probability that some people will be saved and some probability that no people will be saved.

Note that (2a) includes the possibility of *no* people being saved. Crucially, since an outcome in which *some* people are saved for sure is preferable to the possibility of *no* people being saved, FTT predicts that recipients will tend to prefer the sure option (Program A) in the 'be saved' framing condition. This is in line with what is observed in Tversky and Kahneman's experiment (and many others since).

The converse story can be told with respect to the 'die' condition. In this condition, the contrast is between 'some' people dying (under Program C), and 'no' people dying (under Program D). FTT predicts that the sure option in the 'die' condition is represented as in (3a):

- (3a) If Program C is adopted, some people will die.

Meanwhile, the risky option in the 'die' condition is held to be represented as in (4a):

- (4a) If Program D is adopted, there is some probability that nobody will die and some probability that some people will die.

In this condition, then, the risky option is preferable to the sure option; the possibility of *no* people dying is preferable to the certainty that *some* people will die. Thus, FTT predicts that people will tend to prefer the risky option (Program D) in the 'die' condition. Again, this is in line with observed shifts in preferences. Thus, FTT can potentially explain the framing effect induced by the Disease Problem (and similar risky-choice scenarios).

I now turn to the semantics of the gist-like representation, 'some', and begin to discuss the implications of this for FTT's explanation of framing effects.

3 The Semantics of ‘Some’

The proponents of fuzzy-trace theory do not provide an explicit semantics for their ‘gist-like’ representation ‘some’. In particular, they do not say whether it is supposed to have a lower-bounded semantics (‘some and possibly all’) or a bilateral semantics (‘some but not all’). This is unfortunate because it is well established that the linguistic expression ‘some’ is naturally interpreted lower-boundedly in some contexts and bilaterally in others. Take, for example, sentence (5):

- (5) Every student who answers some questions correctly should feel proud.

In (5), ‘some’ is most naturally read as ‘some and possibly all’ (provided there is no special focus on the word ‘some’). Students who answer *all* of the questions correctly should also be proud of themselves – perhaps especially so! In contrast, consider (6):

- (6) This is quite a difficult exam, so some questions can be answered with the aid of a calculator.

In (6) ‘some’ is most naturally read as ‘some but not all’, i.e. there are still some questions that must be answered *without* a calculator. This contextual variation in the interpretation of ‘some’ has led to intense debate in philosophy and linguistics about the expression’s semantic and pragmatic properties: see, for example, Chierchia et al. (2012) and De Neys and Schaeken (2007).

Without getting into the details of that debate here, it is nevertheless reasonable to ask what assumption the fuzzy-trace theorists are making when they claim that numbers get represented as ‘some’: are we to suppose that people are representing a lower-bounded quantity – *some and possibly all* – or are they representing a bilateral quantity – *some but not all*?

On closer inspection, it becomes clear that FTT is, at least implicitly, committed to a *lower-bounded* semantics for ‘some’, denoting *some and possibly all*. This is the only possibility that is consistent with their analysis of the risky options in the Disease Problem. They claim there that the possibility of *all* 600 people being saved (in the ‘be saved’ condition) is represented as a possibility of *some* people being saved. Likewise, the possibility of *all* 600 people dying (in the ‘die’ condition) is held to be represented as a possibility of *some* people dying. Those interpretations would be impossible if ‘some’ meant ‘some *but not all*’.⁷

An implication of this conceptual clarification is the following: insofar as they are represented in gist-like ways, the sure options (Program A and Program C) are *not* represented as being extensionally equivalent. This is in tension with FTT’s explicit claim that “decision makers show valence-dependent preferences despite perceiving options as extensionally equivalent” (Chick et al. 2016, p. 239).

⁷ If the fuzzy-trace theorists were to revise their analysis of these options, so that the possibilities were represented as ones in which ‘all’ people will be saved or will die, then it would no longer be clear what responses the theory would predict. In the ‘be saved’ frame, participants would face a choice between ‘some’ people being saved and either ‘all’ or ‘none’ being saved. In the ‘die’ frame, the choice would be between ‘some’ people dying and either ‘all’ or ‘none’ dying. It is no longer obvious which of these is preferable.

Instead, in the ‘be saved’ condition, FTT is committed to the sure option being represented as in (1b):

- (1b) If Program A is adopted, some and possibly all people will be saved.

Meanwhile, in the ‘die’ condition, the sure option is represented as in (3b):

- (3b) If Program C is adopted, some and possibly all people will die.

Given these representations, the sure options are no longer extensionally equivalent. The sure option in the ‘be saved’ condition is consistent with the possibility of everyone being saved (while ruling out the possibility of everyone dying). In contrast, the sure option in the ‘die’ condition is consistent with everyone *dying* (while ruling out the possibility of everyone being *saved*). At the very least, then, FTT is committed to there being an important sense in which decision-makers perceive – and represent – alternatively framed options as extensionally *inequivalent*. This is necessary for its account of framing effects to be internally consistent.

Could the inequivalence also explain why the sure option is evaluated *more favorably* in the ‘be saved’ condition than in ‘die’ condition? As we will see in the next section, a proposal along these lines has been put forward elsewhere in the framing literature. However, the fuzzy-trace theorists have argued against this ‘lower bounding hypothesis’, along with its promise to give a *rational* explanation of framing effects.⁸ In Section 5, I will argue that, although the fuzzy-trace theorists are right to reject a strong version of the hypothesis, they are already implicitly committed to a modest version of it. Importantly, then, FTT too offers an explanation of framing effects that can partially absolve subjects of behaving irrationally.

4 The Lower Bounding Hypothesis

According to the ‘lower bounding hypothesis’, many recipients interpret the number expressions in the Disease Problem as denoting minimum values. Thus, ‘200’ in the sure option of the ‘be saved’ condition is interpreted as denoting *at least 200*, as in (1c) below:

- (1c) If Program A is adopted, at least 200 people will be saved.

Meanwhile ‘400’ in the sure option of the ‘die’ condition is interpreted as denoting *at least 400*, as in (3c):

- (3c) If Program C is adopted, at least 400 people will die.

As before, under these interpretations, Program A could potentially result in more people being saved than Program C (indeed, it could potentially result in *everyone* being saved, whereas Program C guarantees that 400, at least, will die). This asymmetry potentially

⁸ For further critical discussion of the extent to which the lower bounding account is a genuinely rationalizing one, and some deeper linguistic and philosophical issues this raises, see Fisher (2020).

stands to explain the observed shift in preferences, towards the sure option in the ‘be saved’ condition, and away from the sure option in the ‘die’ condition.

Mandel (2014) is the first to conduct a direct empirical investigation of the lower bounding hypothesis.⁹ Mandel’s results suggest that the framing effect induced by the Disease Problem does, in fact, depend on participants interpreting the number expressions lower-boundedly. First, Mandel fails to find statistically significant framing effects when participants are encouraged to form bilateral (‘exactly’) interpretations of number expressions. Specifically, when participants are presented with modified formulations of the sure options, as in (1d) and (3d), no statistically significant framing effects are observed.¹⁰

- (1d) If Program A is adopted, exactly 200 people will be saved.
- (3d) If Program C is adopted, exactly 400 people will die.

Conversely, Mandel finds that framing effects are *amplified* by encouraging participants to form lower-bounded interpretations. When the sure options include the modifier ‘at least’ (as in (1c) and (3c) above), a large framing effect is observed – larger than for the classic version of the Disease Problem, which contains no modifiers of the number expressions.

Finally, in Mandel’s third experiment, participants are asked whether they interpreted the number expressions in the classic Disease Problem as having ‘at most’, ‘exactly’, or ‘at least’ meanings. Crucially, framing effects are only evident among those who report lower-bounded, ‘at least’, interpretations. No statistically significant framing effect is observed among those who report bilateral, ‘exactly’, interpretations.

In sum, Mandel’s results suggest that the framing effect generated by the Disease Problem is driven (at least in part) by participants interpreting the number expressions in the sure options as denoting lower-bounded quantities.

Mandel’s results are striking but they have been challenged, including by the fuzzy-trace theorists themselves (who classify the lower-bounding hypothesis as a ‘linguistic ambiguity’ account of framing effects – see Chick et al. 2016).¹¹ Chick et al. (2016)

⁹ This is despite the idea being in circulation since the early days of framing research, going back at least as far as Macdonald (1986).

¹⁰ The first experiment tests this using a within-subjects design. The second experiment uses a between-subjects design. In Mandel’s second and third experiments, the following, modified, risky-choice scenario is used:

- In a war-torn region, the lives of 600 stranded people are at stake. Two response plans with the following outcomes have been proposed. Assume that the estimates provided are accurate.

This scenario, taken from Mandel (2001), is intended to make it more plausible that the number of lives under threat could be accurately forecast. The options are then phrased similarly to those in the Disease Problem, except that: (i) the sure options include the phrase ‘it is certain that’ before the number expressions; (ii) the risky options remove 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 is switched; and (iv) each of the options refer to ‘Plans’ rather than ‘Programs’. Additionally, Mandel uses the following wording in the question to participants: ‘Which of the two plans would you choose – A or B?’ I think it is 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 are observed in the other conditions that use the same scenario).

¹¹ The ‘ambiguity’ label reflects the fact that the under-specification of the sure options makes lower-bounded interpretations available. The sure options fail to specify the fate of all 600 people expected to be killed, focusing only on a subset of them. That makes it conceivable that *more than 200* people could end up being saved (in the ‘be saved’ condition), or that *more than 400* could end up dying (in the ‘die’ condition).

demonstrate that statistically significant framing effects *can* arise, even among participants who adopt bilateral interpretations of number expressions. In their second experiment, participants are trained to interpret number expressions as denoting exact quantities, which will not increase or decrease. ‘Ambiguity questionnaires’ are administered before and after the presentation of the Disease Problem. These questionnaires are designed to check whether participants are successfully adopting ‘exactly’ interpretations (and those who are not are removed from the analysis). Importantly, the researchers still observe a framing effect, even when participants interpret the number expressions as denoting exact quantities. This suggests that lower-bounded interpretations are at least not *necessary* for framing effects to arise.¹² Correspondingly, Chick et al. reject a strong version of the lower bounding hypothesis, according to which framing effects depend *only* on lower-bounded representations of quantities.¹³ I will argue in the next section that FTT is nevertheless implicitly committed to a modest version of the lower bounding hypothesis.

5 The Case for Lower Bounding

Although lower-bounded interpretations are unlikely to be *necessary* for framing effects to arise, it remains possible that they are a contributory factor. The fuzzy-trace theorists explicitly acknowledge this possibility but do not pursue it. For example, on one hand they accept that Mandel’s inclusion of ‘exactly’ in the sure options of the Disease Problem attenuates the framing effect, and that “this finding supports the hypothesis that some people do show framing effects because of ambiguity” (Chick et al. 2016, p. 251). On the other hand, they cast doubt on the validity of Mandel’s experiment, writing:

Our disambiguation was more thorough than that provided by Mandel, as we also provided pretask instructions (with worked examples) indicating how to interpret unstated information in both the sure and risky options. We verified that subjects interpreted unstated information as complementary to stated information by quizzing them both before and after they completed the framing problems. By documenting subjects’ interpretations of the options before they made framing decisions, we avoided the demand effect that might have influenced the results reported by Mandel (2014). (Chick et al. 2016, p. 251)

Similarly, while Chick et al. recognize that their own results “do not rule out the possibility that linguistic ambiguity explains some portion of the framing effects reported in the literature” (Chick et al. 2016, p. 252), they end up drawing the following stronger conclusion:

¹² See also Simmons and Nelson (2013), who fail to replicate the second experiment of Mandel (2014). Note, though, that their findings are not necessarily inconsistent with Mandel’s. Indeed, Mandel (2020) responds to Simmons and Nelson, pointing out that both sets of data points fall within a single confidence interval (and also noting a puzzling feature of Simmons and Nelson’s data, which appears to show participants failing to distinguish between the meanings of ‘at least’ and ‘exactly’).

¹³ It is unclear whether anyone actually holds this strong view. It is evident that Mandel, at least, does not: see Mandel (2020) and Tombu and Mandel (2015).

Our results...support the hypothesis that risky choice framing effects are caused by gist-based categorical contrasts between some and none...rather than by interpreting incompletely specified information as ambiguous with respect to the value of the options. (Chick et al. 2016, p. 252)

This is surely too strong. First, as noted above, while lower-bounded interpretations of number expressions may be unable to explain framing effects in their entirety, these interpretations could still explain framing effects in part. Second, I believe FTT should not only *allow* for this possibility but must positively *endorse* it if the account is to be internally consistent. As discussed in Section 3, FTT itself predicts that decision-makers represent number expressions lower-boundedly (as ‘some and possibly all’). Once it is acknowledged that the sure options are represented inequivalently, it is difficult to see why decision-makers should be insensitive to the implication that more people could be saved under Program A than Program C. Certainly, FTT provides no reason to think people should be insensitive to this information. On the contrary, in providing a gloss of Mandel’s findings, Chick et al. say:

Adding “at least” to the sure option makes the value of the sure option higher than that of the risky option in the gain frame, and lower than that of the risky option in the loss frame. Naturally, subjects choose the option with the higher value, resulting in most subjects choosing the sure option in the gain frame and the risky option in the loss frame. (Chick et al. 2016, p. 251)

By the same logic, subjects working with the gist-like representation ‘some and possibly all’ have reason to choose the sure option in the ‘be saved’ condition and the risky option in the ‘die’ condition: in the ‘be saved’ condition, the sure option has a higher value than the risky option; in the ‘die’ condition it has a lower value.

So, once we accept that the number expressions ‘200’ and ‘400’ are represented as lower-bounded quantities (whether as ‘at least 200’ and ‘at least 400’ or merely as ‘some and possibly all’) preferences *should* shift across framing conditions, even if lower bounding is not the *only* factor underlying such a shift. Accordingly, FTT must endorse a modest version of the lower bounding hypothesis, whereby lower-bounded interpretations of number expressions account for at least *part* of the framing effect produced by the Disease Problem.

As discussed, this has important implications for the theory’s other commitments. Insofar as people represent the quantities in the Disease Problem lower-boundedly, they perceive the sure options to be inequivalent. This is in tension with FTT’s explicit claim that people perceive alternatively framed options as equivalent. Importantly, that explicit claim underpins FTT’s ability to maintain that framing effects are irrational. As Chick et al. write of the debate about ambiguity:

This debate has implications for major theories of risky choice. If the ambiguity hypothesis is correct, then framing effects do not represent irrational behavior, after all. This would challenge prospect theory, extended prospect theory, and fuzzy-trace theory accounts of framing effects, all of which assume that decision

makers show valence-dependent preferences despite perceiving options as extensionally equivalent. (Chick et al. 2016, p. 239)

In sum, FTT is incoherent as it currently stands. Making sense of it requires dropping, or refining, the assumption that decision-makers perceive alternatively framed options as extensionally equivalent; and the related claim that framing effects are irrational. In fact, FTT offers an explanation of framing effects that can, at least in part, absolve subjects of behaving irrationally.

6 The Role of Fuzzy Traces

The preceding analysis raises the question of which factor (or factors) accounts for the remainder of the framing effect, i.e. that portion which is not explained by the inequivalence brought about by lower-bounded representations of quantities. Specifically, could the gap be plugged by the core claim of the fuzzy trace account, which appeals to a categorical contrast between 'some' and 'none'? A complete answer to that question lies beyond the scope of this paper, the aim of which has been to show why fuzzy-trace theorists should endorse a modest version of the lower bounding hypothesis. However, in this section I provide some remarks to suggest that support for the core claim of the fuzzy trace account is relatively limited.

The first point concerns FTT's claim that number expressions are preferentially interpreted as 'some' in all kinds of numerical reasoning task. It is worth noting, then, the existence of psycholinguistic research that provides prima facie counterevidence to this claim. In a study by Panizza et al. (2015), it is found that participants resolve problems differently when the number expression 'two' is substituted for 'some'. In principle, though, it seems possible for participants to have completed the experimental task by representing 'two' in exactly the same way as 'some'. This suggests that numbers are not always preferentially represented as 'some'.¹⁴

Second, it seems possible that the unexplained portion of the framing effect could be accounted for equally well by (one or more of) the competing explanations on the market. On the one hand, it should be acknowledged that FTT does provide a relatively neat explanation of some of the framing data. In particular, several studies have tested 'truncations' of the risky options, removing one or other of the clauses, or 'complements', therein. Recall that FTT's explanation of framing effects depends on the contrast between the sure option and the 'zero complement' of the risky option, i.e. the clause concerning the probability of *no* people being saved (in the 'be saved' frame), or the probability of *no* people dying (in the 'die' frame). FTT denies that framing effects depend on the other 'non-zero complement' (which concerns the probability of all 600 people being saved, or all 600 people dying). This is because the gist-like representation of that clause renders it categorically equivalent to the representation of the sure option in each condition.

¹⁴ A natural response for fuzzy-trace theorists here would be to argue that certain elements of Panizza et al.'s experimental set-up prompted subjects to work with an uncharacteristically precise representation of the quantity involved, perhaps because two is such a *small* number. Alternatively, they might argue that the mental representation of the word 'some' can differ from the gist representation 'some'. However, these suggestions merely highlight the need for further empirical investigation of how numbers are represented in a wider range of conditions.

FTT thus predicts that framing effects will only be observed when the zero complements are included in the risky option. Moreover, it predicts that framing effects will be strongest when the zero complements are presented on their own, without the non-zero complements, since that makes the contrast between ‘some’ and ‘none’ even starker. These predictions have been borne out empirically, providing some prima facie support for FTT (Chick et al. 2016; Kühberger and Tanner 2010; Reyna and Brainerd 2011; Reyna et al. 2014). In particular, FTT seems to do a better job of explaining the truncation data than prospect theory (which predicts that framing effects will still be observed when only the non-zero complements are presented).

It is less clear, though, how FTT compares against various other hypotheses. For example, Tombu and Mandel (2015) put forward an ‘Explicated Valence Account’, according to which framing effects depend on the overall positive or negative valence of each option’s description. In the ‘be saved’ condition, the ‘explicated valence’ of the sure option is entirely positive, since it states only how many people will be saved. However, the explicated valence of the risky option is mixed, since it includes the probability of all 600 people being saved (positive) and the probability of no people being saved (negative). Conversely, in the ‘die’ frame the explicated valence of the sure option is entirely negative, whereas the explicated valence of the risky option remains mixed. Tombu and Mandel propose that decision-makers tend to maximize positive explicated valence. This could help explain why they shift towards preferring the sure option in the ‘be saved’ condition but the risky option in the ‘die’ condition. The Explicated Valence Account could also potentially account for the truncation data described above (Tombu and Mandel 2015).

Other promising approaches include some of those Chick et al. classify under their broad ‘linguistic ambiguity’ umbrella. For example, Kühberger and Gradl (2013) suggest that risky-choice framing effects are driven by the *sure* options in the classic version of the Disease Problem, not by the risky options. If they are correct about this, the standard pattern of responses may be analyzable as a straightforward ‘attribute framing effect’ under the typology developed by Levin et al. (1998). As such, it may be amenable to a range of other explanations, such as the associationist account put forward by Levin and colleagues (Levin 1987; Levin and Gaeth 1988; Levin et al. 1986; Levin et al. 1985; Levin et al. 1998) or the ‘information leakage’ account developed by McKenzie and Sher (Leong et al. 2017; McKenzie and Nelson 2003; Sher and McKenzie 2006, 2008, 2011).

As stated above, these cursory remarks certainly do not rule out the possibility that FTT’s ‘gist-like’ representations contribute to framing effects. However, I suggest that showing this to be the case will require significant further work, both to provide independent evidence that number expressions are preferentially interpreted as ‘some and possibly all’; and to assess FTT in relation to other extant accounts of framing effects. Summarizing the discussion in this section, then, the core claim of the fuzzy trace account – that the contrast between ‘some’ and ‘none’ contributes to framing effects – currently rests on rather weaker foundations than the modest version of the lower bounding account to which FTT is also implicitly committed.

7 Conclusion

I have argued that FTT is implicitly committed to a moderate version of the lower bounding hypotheses, according to which lower-bounded representations of quantities

contribute to framing effects (without explaining them entirely). In particular, fuzzy-trace theory predicts that the number expressions in the classic Disease Problem are represented as denoting lower-bounded quantities (namely, *some and possibly all*). As a result, the sure options (Programs A and C) end up being represented as extensionally inequivalent; and this can be expected to contribute to framing effects. This shows FTT to be incoherent as it currently stands. Making sense of it requires dropping, or refining, the claim that decision-makers perceive alternatively framed options as extensionally equivalent; and the related claim that framing effects are irrational. Finally, I noted that much more work is required to show that FTT can account for the portion of the framing effect that is not explained by lower bounding. I suggested that this work might proceed in two directions: first, to demonstrate that numbers are represented as 'some and possibly all' across a wider range of numerical reasoning tasks; and, second, to show that FTT explains the relevant framing data better than rival hypotheses on the market. In the meantime, let us at least agree on having explained *some* of the framing effect!

Acknowledgements This work was supported by the Arts and Humanities Research Council funded South, West and Wales Doctoral Training Partnership (grant number AH/L503939/1). I am grateful to Emma Borg, Nat Hansen, David Mandel, and an anonymous reviewer for this journal, for their extremely helpful comments on earlier versions of the paper.

Funding Open access funding provided by University of Vienna.

Declarations This work was supported by the Arts and Humanities Research Council funded South, West and Wales Doctoral Training Partnership (grant number AH/L503939/1).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Brainerd, C.J., and J. Kingma. 1984. Do children have to remember to reason? A fuzzy-trace theory of transitivity development. *Developmental Review* 4: 311–377. [https://doi.org/10.1016/0273-2297\(84\)90021-2](https://doi.org/10.1016/0273-2297(84)90021-2).
- Brainerd, C.J., and J. Kingma. 1985. On the independence of short-term memory and working memory in cognitive development. *Cognitive Psychology* 17: 210–247. [https://doi.org/10.1016/0010-0285\(85\)90008-8](https://doi.org/10.1016/0010-0285(85)90008-8).
- Broniatowski, D.A., and V.F. Reyna. 2018. A formal model of fuzzy-trace theory: Variations on framing effects and the Allais paradox. *Decision* 5 (4): 205–252. <https://doi.org/10.1037/dec0000083>.
- Chick, C.F., V.F. Reyna, 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* 42 (2): 238–256. <https://doi.org/10.1037/xlm0000158>.

- Chierchia, G., D. Fox, and B. Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In *Semantics: An international handbook of natural language meaning*, ed. C. Maienborn, K. von Heusinger, and P. Portner, vol. 3, 2297–2332. Berlin: Mouton de Gruyter.
- De Neys, W., and Schaeken, W. (2007). When people are more logical under cognitive load: dual task impact on scalar implicature. *Experimental Psychology* 54 (2): 128–133.
- Fisher, S. (2020). Rationalising framing effects: At least one task for empirically informed philosophy. *Crítica. Revista Hispanoamericana De Filosofía* 52 (156): 5–30. <https://doi.org/10.22201/iifs.18704905e.2020.1221>.
- Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–291. <https://doi.org/10.2307/1914185>.
- Kühberger, A. 1998. The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes* 75 (1): 23–55. <https://doi.org/10.1006/obhd.1998.2781>.
- Kühberger, A., and P. Grادل. 2013. Choice, rating, and ranking: Framing effects with different response modes. *Journal of Behavioral Decision Making* 26 (2): 109–117. <https://doi.org/10.1002/bdm.764>.
- 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* 23 (3): 314–329. <https://doi.org/10.1002/bdm.656>.
- Leong, L.M., C.R.M. McKenzie, S. Sher, and J. Müller-Trede. 2017. The role of inference in attribute framing effects. *Journal of Behavioral Decision Making* 30 (5): 1147–1156. <https://doi.org/10.1002/bdm.2030>.
- Levin, I.P. 1987. Associative effects of information framing. *Bulletin of the Psychonomic Society* 25 (2): 85–86. <https://doi.org/10.3758/BF03330291>.
- Levin, I.P., and G.J. Gaeth. 1988. How consumers are affected by the framing of attribute info. *Journal of Consumer Research* 15 (3): 374. <https://doi.org/10.1086/209174>.
- Levin, I.P., R.D. Johnson, P.J. Deldin, L.M. Carstens, L.J. Cressley, and C.R. Davis. 1986. Framing effects in decisions with completely and incompletely described alternatives. *Organizational Behavior and Human Decision Processes* 38: 48–64. [https://doi.org/10.1016/0749-5978\(86\)90025-7](https://doi.org/10.1016/0749-5978(86)90025-7).
- Levin, I.P., R.D. Johnson, C.P. Russo, and P.J. Deldin. 1985. Framing effects in judgment tasks with varying amounts of information. *Organizational Behavior and Human Decision Processes* 36: 362–377. [https://doi.org/10.1016/0749-5978\(85\)90005-6](https://doi.org/10.1016/0749-5978(85)90005-6).
- 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* 76 (2): 149–188. <https://doi.org/10.1006/obhd.1998.2804>.
- Macdonald, R.R. 1986. Credible conceptions and implausible probabilities. *The British Journal of Mathematical and Statistical Psychology* 39 (1): 15–27. <https://doi.org/10.1111/j.2044-8317.1986.tb00842.x>.
- Mandel, D.R. 2001. Gain-loss framing and choice: Separating outcome formulations from descriptor formulations. *Organizational Behavior and Human Decision Processes* 85 (1): 56–76. <https://doi.org/10.1006/obhd.2000.2932>.
- Mandel, D. R. (2014). Do framing effects reveal irrational choice? *Journal of experimental psychology. General*, 143(3), 1185–1198. doi:<https://doi.org/10.1037/a0034207>.
- Mandel, D. R. (2020). Framing effects, replications, and scientific inference: Reply to commentary by Simmons and Nelson (2013) on Mandel (2014). Doi:<https://doi.org/10.31234/osf.io/34jeg>.
- Maule, A.J. 1989. Positive and negative decision frames: A verbal protocol analysis of the Asian disease problem of Tversky and Kahneman. In *Process and structure in human decision making*, ed. H. Montgomery and O. Svenson. Chichester: Wiley.
- McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review: A Journal of the Psychonomic Society, Inc.* 10(3), 596. doi: <https://doi.org/10.3758/BF03196520>, 10, 596, 602.
- Panizza, D., Y.T. Huang, G. Chierchia, and J. Snedeker. 2015. Relevance of polarity for the online interpretation of scalar terms. *Semantics and Linguistic Theory* 360. <https://doi.org/10.3765/salt.v0i0.2530>.
- Reyna, V.F., and C.J. Brainerd. 1991. Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making* 4 (4): 249–262. <https://doi.org/10.1002/bdm.3960040403>.
- Reyna, V.F., and C.J. Brainerd. 2011. Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review* 31 (2): 180–206. <https://doi.org/10.1016/j.dr.2011.07.004>.
- Reyna, V.F., C.F. Chick, J.C. Corbin, and A.N. Hsia. 2014. Developmental reversals in risky decision making: Intelligence agents show larger decision biases than college students. *Psychological Science* 25 (1): 76–84. <https://doi.org/10.1177/0956797613497022>.

- Sher, S., and C.R.M. McKenzie. 2006. Information leakage from logically equivalent frames. *Cognition* 101 (3): 467–494. <https://doi.org/10.1016/j.cognition.2005.11.001>.
- Sher, S., and C.R.M. McKenzie. 2008. Framing effects and rationality. In *The probabilistic mind: Prospects for Bayesian cognitive science*, ed. N. Chater and M. Oaksford, 79–96. Oxford: Oxford University Press.
- Sher, S., & McKenzie, C. R. M. (2011). Levels of information: A framing hierarchy. In G. Keren (Ed.), *Perspectives on framing* (pp. 35-63). Abingdon, Oxfordshire; New York: Psychology press.
- Simmons, J., & Nelson, L. (2013). "Exactly": The most famous framing effect is robust to precise wording. Retrieved from <http://datacolada.org/11>.
- 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* 28 (5): 464–476. <https://doi.org/10.1002/bdm.1863>.
- Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (4481): 453–458. <https://doi.org/10.1126/science.7455683>.
- Wang, X.T. 1996. Framing effects: Dynamics and task domains. *Organizational Behavior and Human Decision Processes* 68: 145–157. <https://doi.org/10.1006/obhd.1996.0095>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.