Framing effects as a semantic puzzle: Putting the alignment-assumption account to a test

Berry CLAUS — *Humboldt-Universität zu Berlin*

Abstract. Framing effects are effects of linguistic variation (e.g. stating the amount of lives saved vs. lives lost) on judgments and decisions. This paper deals with a semantic-pragmatic account of framing effects as offered by Geurts' (2013) alignment-assumption account. The account radically differs from extant accounts by explaining framing effects in terms of counterfactual alternatives and alignment of scales. I report two experiments that tested predictions derived from the alignment-assumption account and that related to the effect of upward- vs. downward entailing comparative quantifiers. The results provide preliminary experimental support for the alignment-assumption account and pose challenges for other accounts of framing effects.

Keywords: framing effects, comparative quantifiers, evaluative predicates, entailment, scales.

1. Introduction

The term *framing effect* refers to the well-established finding that choices and judgments are systematically affected by varying the description of options and states of affairs. In the first demonstration of a framing effect (Tversky and Kahneman, 1981), participants were presented with the scenario of an imminent outbreak of a deadly disease expected to kill 600 people (see (1)). They were asked to choose between two alternative programs to combat the disease, with one program having a sure outcome (e.g. Program A1) and the other one a risky outcome (e.g. Program A2). The outcomes were presented either in terms of number of lives saved, as in (1), or in terms of number of lives lost, as in (2).

(1) Imagine that 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 that the exact scientific estimates of the consequences of the programs are as follows:

If Program A1 is adopted, 200 people will be saved.
If Program A2 is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Which of the two programs would you favour?

(2) If Program B1 is adopted, 400 people will die.
If Program B2 is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

The different framing of the options had a strong effect on participants' choices. In the *survive*-frame condition, the majority (72%) of the participants chose the program with the sure outcome, whereas in the *die*-frame condition, it was chosen only by the minority (22%) of the participants.

Proceedings of Sinn und Bedeutung 23, vol. 1, pp. 249–266. Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès).

Framing effects[1] are a robust finding; they have been demonstrated in numerous experimental studies (albeit the effects tend to be smaller than in the study by Tversky and Kahneman). Their source, however, is a matter of controversy. Accounts of framing effects range from questioning the truth-conditional equivalence of differently framed descriptions to attributing them to irrationality. In between are accounts that explain framing effects in terms of implicitly conveyed information. Considering the very basis of framing effects, they seem to cry out for a semantic-pragmatic account. Yet, framing effects have received only limited attention in the linguistic literature. A notable exception is the alignment-assumption account of Geurts (2009; 2013), which opens up a new perspective on framing effects by explaining them in matters of counterfactual alternatives and alignment of scales.

In what follows, I first give an overview of accounts of framing effects. Section 3 gives a description of the alignment-assumption account (Geurts, 2009; 2013). In Section 4, two experiments are reported that were designed to test predictions derived from the alignment-assumption account. The findings of the experiments are discussed within the framework of the alignment-assumption account as well as with regard to their implications for other accounts of framing effects. Section 5 concludes the paper.

2. Accounts of framing effects

Accounts of framing effects can be divided in three groups. They differ in their underlying assumptions about the equivalence of two differently framed descriptions. First, there are accounts that do not question the equivalence of the two description variants and explain framing effects in terms of representational differences. According to the second group of accounts, framing effects emerge because the descriptions in the two frames are not necessarily truth-conditionally equivalent. Third, there are accounts in which it is assumed that the two description variants are not information equivalent and in which the framing effect is attributed to a difference in implicit information between the two frames.

2.1. Accounts of group 1: representational difference between the two frames

Tversky and Kahneman (1981) interpreted their finding of a framing effect in terms of prospect theory, i.e., as an effect of framing on risk propensity. Prospect theory was developed by Kahneman and Tversky (1979) as a descriptive account of decision behavior under risk. In a nutshell, the central proposal of prospect theory is that choices involving gains tend to trigger risk aversion whereas choices involving losses tend to trigger risk seeking. In particular, it is presumed that the outcomes of choice options are perceived as gains or losses relative to

[1] In the literature, different types of framing effects are distinguished (Levin, Schneider and Gaeth, 1998). The framing effect in the study by Tversky and Kahneman is an example of *risky-choice framing*. A further main type of framing is *attribute framing*. In attribute framing tasks, participants do not have to choose between options, but rather are presented with an object to which a particular attribute is ascribed, either with a positive or a negative frame (e.g. The yoghurt is 90% fat free / contains 10% fat). Their task is to evaluate the object or, for example, to indicate whether they would buy it. Both types of framing effects are empirically well-established. In what follows, I will focus on risky-choice framing. However, the accounts of framing-effects that are described below can be applied to both types of framing.

a reference outcome, which is assigned a value of zero. Alternative descriptions of choice options may induce different reference outcomes[2]. Regarding the deadly-disease decision problem, it is assumed that the *survive*-frame induces a reference outcome of zero lives saved, such that the outcome of the sure option (200 lives saved) is perceived as a gain. In contrast, the *die*-frame induces a reference outcome of zero lives lost, such that the outcome of the sure option (400 lives lost) is perceived as a loss. In prospect theory, it is further proposed that the value function for gains differs from the value function for losses. The convex function for losses is steeper than the concave function for gains, such that losses "loom larger" than gains. Uncertain outcomes are assumed to be multiplied by a weight reflecting a nonlinear perception of probabilities, such that high and mid-level probabilities are underweighted more strongly than low probabilities are overweighted. Applied to the intervention programs in the deadly-disease scenario, the risk aversion in the *survive*-frame (see (1)) is due to the value of the gain of 200 lives saved being larger than the weighted value associated with a 1/3 probability of 600 lives saved. On the other hand, the risk seeking in the *die*-frame (see (2)) is due to that the negative value associated with 400 lives lost is smaller than the value associated with a 2/3 probability of 600 lives lost.

According to fuzzy-trace theory (Reyna and Brainerd, 1991; see Setton et al., 2014 for an overview), risky choice framing effects are not driven by fine-grained (subjective) representations of the stated quantities. Rather, people base their choices on coarse "gist" representations of the options' qualitative meaning (as long the gist representations are sufficient to discriminate between options)[3]. This key assumption of fuzzy-trace theory was inspired (cf. Reyna 2012) by the psycholinguistic distinction between verbatim memory and gist memory. A gist representation of the positively framed descriptions of the two options in the deadly-disease scenario (see (1)) is given in (3).

(3) If Program A1 is adopted, some people will be saved.
If Program A2 is adopted, some will be saved or no one will be saved.

In the gist representation in (3), quantities are reduced to dichotomies, i.e., 'some' vs. 'no one', and the probabilities in the risky option are reduced to a disjunction. As a consequence, both options include the outcome of saving some people and the decision between the two programs boils down to a decision between saving some people (Program A1) and saving no one (Program A2), rendering Program A1 as the better option (given that it is better that more people are saved than that less people are saved). (4) illustrates the gist representations of the negatively framed options (see (2)).

(4) If Program B1 is adopted, some people will die.
If Program B2 is adopted, no one will die or some people will die.

[2] The assumptions with regard to the assignment of reference outcomes seem to be largely intuition-based. That is, it is not clearly spelled out which factors determine the reference outcome (Werner and Zank, 2018).

[3] Fuzzy-trace theory assumes that in addition to gist representations, people encode verbatim representations to which they resort when the gist representations of two options do not differ – rendering the account difficult to falsify.

In (4), the decision is reduced to a decision between some people dying (Program B1) and no one dying (Program B2), rendering Program B2, i.e., the risky option, as the better option. In comparison to prospect theory, fuzzy-trace theory offers a more parsimonious account of framing effects. Moreover, findings of studies within the field of fuzzy-trace theory indicate that the representation of numerical information may not be the driving factor behind framing effects. For example, a strong framing effect still emerged when the options were described as in (3) and (4), i.e., when all numerical information was omitted (Reyna and Brainerd, 1991). This finding cannot be readily explained within the framework of prospect theory. Note, however, that the assumed gist representations of the options are somewhat ad hoc. For instance, it is unclear for which independent reasons, the gist representations for *1/3 probability that 600 people will be saved* (see (1)) and *2/3 probability that 600 people will die* (see (2)) are assumed to be 'some will be saved' (see (3)) and 'some will die' (see (4)), respectively, rather than 'all will be saved' and 'all will die'.

2.2. Accounts of group 2: truth-conditional difference between the two frames

The so-called ambiguity hypothesis starts out from the conjecture that the description of the sure option of the deadly-disease scenario is incomplete (e.g., Kühberger, 1995; Kühberger and Tanner, 2010; Mandel, 2001). For example, the descriptions of Program A1 in (1) and Program B1 in (2) leave open the fate of the people who are not mentioned. Of course, it is plausible to assume that participants will infer for Program A1 that all of the 400 unmentioned people will not survive the disease, and for Program B1 that all of the 200 unmentioned people will not die from the disease. With these inferences, the two different frames are truth-conditionally equivalent. However, the incompleteness of the description leaves room for alternative inferences, or rather for wild guesses. For example, people might act on the assumption that some of the 200 people who are not mentioned in the *survive*-frame description of Program A1 (see (1)) will also survive and/or the assumption that some of the 400 people who are not mentioned in the *die*-frame description of Program B1 (see (2)) will also die from the disease. If so, then the truth conditions of the two different frames diverge, as Program A1 and Program B1 involve two different outcomes (200 + n out of 600 lives saved ≠ 400 + n out of 600 lives lost). Along these lines of reasoning, it could be argued that risky-choice framing effects are simply due to the fact that the given frames are not necessarily conceived as being truth-conditional equivalent. Accordingly, framing effects should be eliminated when the outcome of the sure option is fully specified, as in (5). In line with this prediction, no framing effect was found in studies employing full-fledged descriptions of the sure option (Druckman, 2001; Kühberger 1995; Mandel, 2001, 2014)[4].

(5) a. If Program A1 is adopted, then 200 people will survive, and 400 people will not survive.
b. If Program B1 is adopted, then 400 people will die, and 200 people will not die.

[4] Note that this finding is neither in conflict with the fuzzy-trace theory account of framing effects, nor with any of the accounts described below (lower-bound reading, information leakage, argumentative-orientation, alignment assumption) considering that the variants of complete descriptions of the sure option do not differ in the variables considered decisive in these accounts.

A notable variant of the ambiguity hypothesis relates to the interpretation of the numerals that are contained in the descriptions of the sure options. Although there is no generally accepted treatment of numerals, it is uncontroversial that bare numerals can receive different readings (e.g. Breheny, 2008; Carston, 1998; Geurts, 2006; Horn, 1992; Kennedy, 2015). The issue of numeral interpretation has largely been ignored in research on framing effects. Yet, as has been pointed out by Mandel (2014; see also Kühberger, 1995), the readings that participants of framing studies assign to the given numerals are critical with regard to the interpretation of framing effects. Mandel (2014) proposes a lower-bound-reading account and assumes a tendency to assign a lower-bound reading to the numerals in the (not fully specified) sure options. If true, then the positively and negatively framed descriptions are not truth-conditionally equivalent (at least 200 out of 600 lives saved ≠ at least 400 out of 600 lives lost), offering a simple account of framing effects. Mandel (2014) investigated the role of numeral interpretation in a series of experiments. In one of the experiments, participants had to indicate their interpretation of the given numerals by choosing between three alternatives: 'exactly', 'at least', 'at most'. The majority of participants (64%) indicated a lower-bound interpretation of the numeral. Moreover, a standard framing effect obtained only for this subgroup of participants. However, these findings do not provide unequivocal evidence for the lower-bound-reading account, due to two potentially critical methodological aspects. First, the forced choice task did not include an option for indicating an approximate reading of the numerals. This is problematic, considering that it is not unlikely that round numbers such as 200 and 400 are interpreted as approximate estimates (e.g., Krifka, 2009). Second, the numeral interpretation assessment immediately followed the task of choosing between the sure and risky option. Hence, the just-made choice might have influenced participants' indication of the numeral interpretation (e.g., in terms of a justification of the choice)[5]. In two other experiments, Mandel (2014) found that the framing effect was absent, when a precise reading of the numerals was enforced by modifying them with *exactly*. This finding is consistent with the lower-bound-reading account. However, there is a conflicting finding from a study in which a strong framing effect obtained when the numerals were likewise modified with *exactly* (Chick, Reyna and Corbin, 2016).

2.3. Accounts of group 3: implicit-information difference between the two frames

The information-leakage account (McKenzie and Nelson, 2003; Sher and McKenzie, 2006) differs from the preceding accounts in two aspects. Firstly, it neither questions the truth-conditional equivalence of differently framed descriptions nor does it assume that they are fully equivalent. Rather, the key assumption of the information-leakage account is that differently framed descriptions are not information equivalent in the sense that they implicitly convey different information. Secondly, the information leakage account adopts a somehow communicative perspective by assuming that listeners draw inferences as to why the speaker chose a given frame. Thus, the use of a particular frame is assumed to "leak" information on

[5] I am currently conducting an experiment with a numeral-interpretation assessment with four alternatives and with distractor tasks in between the task of choosing between the intervention programs and the assessment of the numeral interpretation. The preliminary results are inconsistent with the lower-bound-reading account. The vast majority of participants indicated that they interpreted the numerals with a precise reading (79%; approximate: 12%; lower bound: 6%; upper bound: 3%). Moreover, a framing effect obtained for all participants as well as for the subgroup of participants who indicated a precise reading.

the speaker's state of mind. More specifically, the speaker's choice of a frame is assumed to be affected by whether and how the to be described state of affairs deviates from a reference point, e.g., from expectations and/or the standard in the given domain. Consider the outcome of a novel medical treatment that could either be described in terms of the survival rate or in terms of the mortality rate. According to the information-leakage account, a speakers' choice between these frames depends on how the outcome of the new medical treatment differs from a reference point, e.g., the outcome of the standard medical treatment, such that she will select the frame that involves an increased rate relative to a reference point. For example, she will use a *survive*-frame as in (6a) rather than a *die*-frame as in (6b), if the survival rate of the standard treatment is less than 50% and vice versa if the standard treatment has a mortality rate less than 50%.

(6) a. 50% of patients with the new treatment survive.
b. 50% of patients with the new treatment die.

According to the information-leakage account, framing effects arise because the two frames convey different implicit information as to how the options differ from the reference point. From the *survive*-frame, it can be inferred that more lives will be saved than with the reference point, whereas from the *die*-frame, it can be inferred that more lives will be lost than with the reference points. Empirical evidence supports the assumption that frame choice depends on how the to be described state of affairs deviates from a reference point (McKenzie and Nelson, 2003; Sher and McKenzie, 2006). Additional findings suggest that the effect of reference point on frame selection goes along with corresponding inferences on the listener's side (McKenzie and Nelson, 2003; Sher and McKenzie, 2006). For example, when being presented with a *survive*-frame as in (6a) participants were more likely to assume that the new medical treatment had a higher survival rate than the old treatment (McKenzie and Nelson, 2003). However, though these findings are consistent with the underlying assumptions of the information-leakage account, they do not provide direct evidence that framing effects emerge from reference-point inferences.

The argumentative-orientation account (Holleman and Pander Maat, 2009) is similar in spirit to the information leakage account as it also explains framing effects in terms of implicit information. However, the two accounts differ, in that the argumentative-orientation account focuses on the role of the communicators' intentions, which are neglected in the information leakage account. Holleman and Pander Maat couch their account in terms of generalized conversational implicatures. They assume that speakers' frame selection is guided by their argumentative orientation, i.e., by the direction of the conclusion the listeners are intended to draw from the utterance. Complementary, listeners make argumentative-orientation inferences. That is, based on the uttered frame, they infer the direction of the conclusions intended by the speaker. [6] For example, framing the outcome of an exam in terms of passing rate rather than failure rate may result in different argumentative-orientation inferences, e.g., with regard to the easiness/difficultness of the exam. The strength of argumentative-orientation inferences is assumed to be affected by the markedness of the chosen frame. It is stronger for marked

[6] A similar view on framing is taken by Mercier and Sperber (2011), who interpret framing effects as support for their argumentative theory.

frames than for unmarked frames. [7] The argumentative-orientation account implies that framing effects arise because the two frames trigger different argumentative-orientation implicatures, e.g. that the conclusion intended by the speaker is either to choose (*survive*-frame) or not to choose the option (*die*-frame). Holleman and Pander Maat report experimental findings that are consistent with the notion that a speaker's argumentative orientation affects which frame she chooses, as well as findings that are in line with the converse notion, that a given frame conveys implicit information on the speaker's intended argumentative conclusion. Yet, these findings do not directly show that the source of framing effects lies in argumentative-orientation implicatures.

3. Alignment-assumption account

Imagine a scenario where 600 people were infected by a potentially deadly disease and 200 of the 600 people survived the disease and the remaining 400 did not survive the disease. To put it in other words, 400 of the 600 people died of the disease and the remaining 200 did not die of the disease. In this scenario, (7a) and (7b) can be considered to be truth-conditionally equivalent. Yet, when they are embedded under the positively evaluative predicate *good* the resulting assertions (8a) and (8b) appear contradictory.

(7) a. Two-hundred people survived the disease.
b. Four-hundred people died of the disease.

(8) a. It's good that 200 people survived the disease.
b. It's good that 400 people died of the disease.

To solve this semantic problem, Geurts (2013) proposes that expressions like *good* trigger the adoption of an alignment assumption. [8] More specifically, *good*(φ) presupposes that φ's alternatives (Alt(φ)) are ordered on a qualitative "goodness" scale. If these alternatives can also be ordered on a quantitative scale in terms of entailment-based strength, then the two scales are assumed to be aligned, as indicated by the formulation of the alignment assumption in (9).[9] For the evaluative statements in (8), the alignment assumption implies the inferences in (10), which clearly contradict each other. Thus, Geurts' notion of an alignment assumption explains why one cannot consistently endorse both (8a) and (8b).

(9) $\forall\psi, \psi' \in \text{Alt}(\varphi), \psi > \psi' \rightarrow \psi \gg \psi'$ >: 'stronger than'; ≫: 'better than'

(10) a. $n + 1$ survived ≫ n survived
b. $n + 1$ died ≫ n died

[7] Holleman and Pander Maat assume that markedness of a frame is determined by its valence (positive = unmarked; negative = marked) and its goal salience in a given scenario (goal consistent = unmarked (e.g., full in a filling scenario; empty in an emptying scenario); goal inconsistent = marked (e.g., full in an emptying scenario; empty in a filling scenario)).

[8] He suggests that "the core meaning of 'good' is something like the following: 'It's good that φ' means that φ ranks sufficiently highly on the relevant qualitative scale which orders Alt(φ)" (Geurts, 2013: p. 10).

[9] In Geurts (2009), two weaker versions of the alignment assumption are discussed.

The notion of an alignment assumption also explains why the two evaluative statements in (11), in which the numerals are modified with the comparative modifiers *more than* and *fewer than*, do not seem contradictory. The two inferences in (12) are compatible. (11b), with the downward-entailing *fewer than*, differs from (8a), (8b), and (11a), in that its alternatives become stronger with decreasing rather than increasing number.

(11) a. It's good that more than 200 people survived the disease.
b. It's good that fewer than 400 people died of the disease.

(12) a. more than $n + 1$ survived ≫ more than n survived
b. fewer than n died ≫ fewer than $n + 1$ died

Geurts applies his alignment-assumption solution of a semantic problem to framing effects by the following line of reasoning: Judgments and decisions do not only reflect the evaluation of a particular option, but also the evaluation of its counterfactual alternatives. The choice task in framing studies involves an evaluation process which can be expected to be affected by the alignment assumption. A person who is presented with the *survive*-frame of the sure option (Program A1) will evaluate the option as positive to some degree (and would rate an option with a higher survival rate as more positive). However, when instead being presented with the *die*-frame of the sure option (Program B1), the evaluation will tend to be negative, as a positive evaluation would involve the alignment-assumption inference that a higher number of lives lost would be more positive. Thus, according to the alignment-assumption account, the framing effect is driven by the evaluation of the sure option. Indeed, there is experimental evidence that the framing effect is not due to having to choose between a sure and a risky option, but rather is due to an attractiveness difference between the two variants of the sure option (Kühberger and Gradl, 2013; Peters and Levin, 2008).

4. Testing two predictions of the alignment assumption account

Geurts' (2013) alignment-assumption account radically differs from other accounts of framing effects. To my knowledge, it is not in conflict with the extant findings. Rather, it opens up a new perspective and leads to novel predictions. In the following, I will address two predictions that relate to the effect of modifying the numerals in the two description versions of the sure option with an upward-entailing comparative quantifier (see (13a)) and a downward-entailing comparative quantifier (see (13b)).

(13) a. If Program A1 is adopted, more than 200 people will survive.
b. If Program B1 is adopted, fewer than 400 people will die.

The alignment-assumption account predicts an interaction effect between the frame (*survive*/*die*) and the way in which the outcome is stated, i.e., with modified numerals as in (13) or with bare numerals as is common in framing studies (see (1) and (2)). The interaction should affect the evaluation of the sure option (see Prediction 1 in (14)) and the pattern of choices of the sure and risky option (see Prediction 2 in (15)).

(14) Prediction 1
The two differently framed descriptions of the sure option in (13) with modified numerals will not result in different evaluative ratings, because the alignment-assumption inferences are compatible. In contrast, with unmodified numerals, the *survive*-variant will be rated more positive than the *die*-variant, because a positive evaluation of the *die*-variant would involve the alignment-assumption inference that a higher number of lives lost would be more positive.

(15) Prediction 2
Different from the standard descriptions of the sure options, i.e., with bare numerals, no framing effect will occur when the numerals in the sure options are modified as in (13).

The predictions were tested in two experiments. The experiments were conducted in German and were implemented as web-based experiments. Experiment 1 addressed Prediction 1 and employed a judgment task; Experiment 2 targeted Prediction 2 and employed a choice task. In both experiments, participants were presented with a cover story (see (16)) that corresponded to the preamble used in the study by Tversky and Kahneman (1981).

(16) Cover story in Experiment 1 and 2
Stellen Sie sich die folgende Situation vor: Die Gesundheitsbehörde einer Kleinstadt bereitet sich auf den Ausbruch einer hochansteckenden Krankheit vor, durch die voraussichtlich 600 Menschen getötet werden. Es gibt eine Vielzahl von unterschiedlichen Programmen zur Bekämpfung der Krankheit, zwischen denen sich die Gesundheitsbehörde entscheiden muss.
'Imagine the following situation: The health authority of a small town is preparing for the outbreak of a highly contagious disease, which is expected to kill 600 people. There are a variety of different programs to combat the disease between which the health authority must decide.'

4.1. Experiment 1

Experiment 1 was designed to test Prediction 1. That is, it investigated whether there is an interaction effect of frame (*survive*/*die*) and numeral (unmodified/modified) on the evaluation of the sure option. The four versions of the sure option in (17) were presented to participants as different intervention programs within the deadly-disease scenario. Their task was to rate each intervention program with regard to how good or bad they consider it to be.

(17) a. Bei Anwendung von Programm A1 werden 200 Menschen überleben.
'If Program A1 is adopted, 200 people will survive.'
b. Bei Anwendung von Programm B1 werden 400 Menschen sterben.
'If Program B1 is adopted, 400 people will die.'
a'. Bei Anwendung von Programm A1 werden mehr als 200 Menschen überleben.
'If Program A1 is adopted, more than 200 people will survive.'
b'. Bei Anwendung von Programm B1 werden weniger als 400 Menschen sterben.
'If Program B1 is adopted, fewer than 400 people will die.'

4.1.1. Method

Participants. The participants of Experiment 1 and Experiment 2 were native speakers of German who were recruited from the student population of Berlin and Potsdam. All participants gave informed consent for participation and participated in exchange for the chance to win € 25 in a raffle. Eighty students (18 to 36 years, M = 22.11; 57 female) participated in Experiment 1. The data of six additional participants were excluded from the analyses because they were not native speakers of German.

Design. Experiment 1 employed a 2x2 within-subject design with the factors NUMERAL (unmodified/modified) and FRAME (*survive/die*).

Materials and Procedure. In the instruction, participants were asked to imagine the scenario of an imminent outbreak of a deadly disease. They were told that there are several intervention programs to combat the disease and that their task was to indicate for each program how good or how bad they judge it to be. There were four experimental trials and 28 filler trials, each with a different intervention program. In all 32 trials, participants were asked to judge the given intervention program on an evaluative rating scale ranging from 1 (very bad) to 7 (very good). In the experimental trials, the description of the intervention programs was based on the sure options of the study by Tversky and Kahneman (1981). The four experimental trials resulted from the 2x2 conditions from the design of Experiment 1 (see (17)). Thus, in two experimental trials, the outcome of the intervention program–in terms of the amount of lives either to be saved or lost–was indicated with an unmodified numeral. In the other two experimental trials, the numerals were modified, i.e., by the upward entailing comparative quantifier *mehr als* ('more than') in the *survive*-frame and by the downward entailing comparative quantifier *weniger als* ('fewer than') in the *die*-frame. In half of the filler trials, the intervention program was described with the *survive*-frame; in the remaining half of the filler trials, it was described with the *die*-frame. The intervention programs in the filler trials differed from those in the experimental trials in the amount of lives to be saved or lost (e.g. *... werden 100 Menschen sterben* [‘… 100 peole will die’] or in aspects of the description, such as different modifying expressions (e.g. *exact* [‘exactly’], *ungefähr* [‘approximately’], *mindestens* [‘at least’]), stating the outcome in terms of percentages (e.g. *... 50% überleben* [‘… 50% survive’] or stating the probability of the outcome (e.g. *... werden mit einer Wahrscheinlichkeit von 2/3 alle sterben* [‘… there is a probability of 2/3 that all will die’]. Participants were randomly assigned to one of 24 different orders of presentation of the experimental and filler trials. The two frame versions of the experimental trials in either of the two modification versions were always intervened with a sequence of 15 other trials. The experimental session lasted approximately 15 minutes.

4.1.1. Results

Table 1 shows the median ratings in the four conditions of Experiment 2. The rating data were analyzed by using a cumulative link mixed model for ordinal data (R package ordinal) with NUMERAL and FRAME as fixed effects and with participants as random effect. For both

fixed factors, sum coding was used (+.5 for 'unmodified' and 'survive'; -.5 for 'modified' and 'die'). The analysis yielded main effects of NUMERAL (b = .41, SE = .15, z = 2.82, p < .01) and FRAME (b = -1.65, SE = .32, z = -5.15, p < .01). These were qualified by a significant interaction of the two factors (b = .69, SE = .21, z = 3.26, p < .001), consistent with Prediction 1. In order to break up the interaction, two separate analyses for the data subsets of the two NUMERAL conditions were conducted. The results of the analyses strengthen the evidence for Prediction 1. In the 'unmodified numeral' condition, the 'survive' version was significantly rated more positive than the 'die' version (b = -1.33, SE = .32, z = -4.19, p < .01). In the 'modified numeral' condition, the ratings did not differ in the two frames (b = -.25, SE = .29, z = -.85, p = .40).

	FRAME	
NUMERAL	'survive'	'die'
umodified	3	2
modified	3	3

Table 1: Median ratings in the four conditions of Experiment 1

4.2. Experiment 2

In Experiment 2, the prediction was tested that the occurrence of a framing effect depends on whether or not the numerals in the sure option are modified (Prediction 2), i.e., that a framing effect will occur with bare numerals, as in the standard description of the sure option, but not when the numerals in the sure options are modified as in (17a') and (17b'). To this end, Experiment 2 employed the standard choice task as in the classical study by Tversky and Kahneman (1981). That is, participants had to choose between two programs, one with a sure outcome and one with a risky outcome. There were four experimental choice tasks, implementing the four conditions addressed in Prediction 2. The four versions of the sure option were the same as in Experiment 1 (see (17)). When the numerals in the sure options were unmodified, the description of the corresponding risky options were German versions of the descriptions used in the study by Tversky and Kahneman (1981), as shown in (18a) and (18b). The sure options with modified numerals were paired with adjusted descriptions of the risky options. Without the adjustment, the expected value of the risky option in the *survive*-frame (200) would have been lower and the expected value of the risky option in the *die*-frame (400) would have been higher than the vague outcome of the sure options (*survive*-frame > 200; *die*-frame: < 400). In order to equal the expected value of the risky options as far as possible with the outcome of the corresponding sure option, the probability statements in the descriptions of the risky options were also modified by comparative quantifiers. That is, for the first mentioned outcome of the risky option, the probability statement was modified with *mehr als* ('more than') and for the second mentioned outcome it was modified with *weniger als* ('less than'), as shown in (18a') and (18b').

4.2.1. Method

Participants Sixty-one students (18 to 36 years, M = 21.79; 45 female) participated in Exper-

iment 2. The data of three additional participants were excluded from the analyses because they were not native speakers of German.
Design Experiment 2 employed a two-factorial within-subject design with the factors NUMERAL (unmodified/modified) and FRAME (survive/die).

(18) a. Bei Anwendung von Programm A2 werden mit einer Wahrscheinlichkeit von 1/3 600 Menschen überleben und mit einer Wahrscheinlichkeit von 2/3 wird niemand überleben.
'If Program A2 is adopted, there is a probability of 1/3 that 600 people will survive, and a probability of 2/3 that no one will survive.'

b. Bei Anwendung von Programm B2 wird mit einer Wahrscheinlichkeit von 1/3 niemand sterben und mit einer Wahrscheinlichkeit von 2/3 werden 600 Menschen sterben.
'If Program B2 is adopted, there is a probability of 1/3 that no one will die, and a probability of 2/3 that 600 people will die.'

a'. Bei Anwendung von Programm A2 werden mit einer Wahrscheinlichkeit von mehr als 1/3 600 Menschen überleben und mit einer Wahrscheinlichkeit von weniger als 2/3 wird niemand überleben.
'If Program A2 is adopted, there is a probability of more than 1/3 that 600 people will survive, and a probability of less than 2/3 that no one will survive.'

b'. Bei Anwendung von Programm B2 wird mit einer Wahrscheinlichkeit von mehr als 1/3 niemand sterben und mit einer Wahrscheinlichkeit von weniger als 2/3 werden 600 Menschen sterben.
'If Program B2 is adopted, there is a probability of more than 1/3 that no one will die, and a probability of less than 2/3 that 600 people will die.'

Materials and Procedure Participants were presented with the same scenario as in Experiment 1 (see (16)) and were given eight choice tasks. In each of the choice tasks, they had to choose between two intervention programs, one with a sure outcome and one with a risky outcome. The eight choice-task trials were intervened by eight distractor trials. In the distractor trials, participants were asked to name four words that begin or end with a given chain of two to three letters. The sequence of trials started with a distractor trial which was followed by an alternating presentation of the eight choice-task trials and the remaining seven distractor trials. Half of the choice-task trials were experimental trials; the other half were filler trials. The four experimental trials corresponded to the four conditions resulting from the design of Experiment 2 (see (17) and (18)). That is, there were two experimental trials with bare numerals, one in the *survive*-frame and one in the *die*-frame. In the two other experimental trials, the numerals were modified by a comparative quantifier, i.e., by *mehr als* ('more than') in the *survive*-frame of the sure option and by *weniger als* ('fewer than') in the *die*-frame of the sure option[10]. In the filler trials, all numerals were modified albeit differently than in the experimental trials (e.g. by *exact* ['exactly']). Two of the filler trials had a *survive*-frame; the other two fillers had a *die*-frame. The four filler choice tasks and the four experimental choice tasks

[10] In the risky options of the trials, the numerals indicating the probabilities of the two outcomes were modified by the two comparative quantifiers in such a way that the (vague) expected value of the risky options matched the outcome of the corresponding sure options (see (18a') and (18b')).

were presented in 24 different orders, which were randomly assigned to the participants. In between the two frame versions of the experimental choice tasks in either of the two modification versions, there were always three other choice tasks and four distractor tasks. The experimental session lasted approximately 15 minutes.

4.2.1. Results

Table 2 shows the proportion of sure-option choices in the four conditions of Experiment 2. The data were analyzed by using a generalized linear mixed model with a binomial logit function and with participants as random factor. There were two fixed factors, NUMERAL and FRAME, both with sum coding (+.5 for 'unmodified' and 'survive'; -.5 for 'modified' and 'die'). There was no main effect of FRAME (b = -.48, SE = .33, z = -1.44, p = .15) and a significant main effect of NUMERAL (b = 1.15, SE = .35, z = 3.29, p < .01). The latter was qualified by a significant interaction of both factors (b = -1.65, SE = .68, z = -2.43, p < .05). To examine the interaction, two separate analyses for the data subsets of the two NUMERAL conditions were conducted. In the 'unmodified numeral' condition, there was a significant effect of FRAME on the choice patterns (b = -1.48, SE = .56, z = -2.64, p < .01), i.e., in the 'survive' condition, the majority of participants (64%) chose the sure option, whereas in the 'die' condition, the majority of participants (57%) chose the risky option. In the 'modified numeral' condition, the choice pattern was not affected by the frame (b = .30, SE = .45, z = .66, p = .51), i.e., in the 'survive' condition as well as in the 'die' condition, most participants favoured the sure option over the risky option (see Table 2). This pattern of results is consistent with Prediction 2. A framing effect obtained for the 'unmodified numeral' versions of the description of the intervention programs. In contrast, for the 'modified numeral' versions, there was no framing effect.

	FRAME	
NUMERAL	survive	die
unmodified	64%	43%
modified	69%	74%

Table 2: Proportion of sure-option choices in the four conditions of Experiment 2

4.3. Discussion

In Experiment 1, no effect of frame (survive/die) on evaluative ratings was found, and in Experiment 2, no framing effect on the choice patterns for the sure vs. risky option occurred, when the numerals indicating the number of lives saved and lives lost were modified with the upward-entailing comparative quantifier *mehr als* and the downward-entailing comparative quantifier *weniger als*, respectively. In contrast, with a standard description of the options, i.e., with unmodified numerals, the *survive*-version of the sure option was rated significantly more positive than the *die*-version (Experiment 1). In addition, there was a significant framing effect with unmodified numerals in that the sure option was preferred in the 'survive' condition and dispreferred in the 'die' condition (Experiment 2).

The interpretation of null effects can be tricky. However, for two reasons it seems unlikely that the null effects obtained with the modified numerals are due to a failure to detect an effect of the frame manipulation when in truth there was one. First, the null result was observed in two experiments. Second, the fact that there were significant effects with unmodified numerals indicates that the materials and methods were sufficiently sensitive to reveal an effect of the manipulated frame.

The present findings provide preliminary experimental evidence for the two predictions that were derived from the alignment-assumption account. Within that account, the findings can be interpreted in terms of alignment-assumption inferences that come along with valence evaluations. Hence, the absence of effects of the frame manipulation with modified numerals can be attributed to that a positive evaluation of the sure option for both frames involves alignment-assumption inferences that conform to human convictions (the more lives saved/the less lives lost the better). [11] By the same token, the presence of effects of the frame manipulation with unmodified numerals can be explained by that the alignment-assumption inferences involved with a positive evaluation of the sure option are in concord with human convictions for the *survive*-frame but are in conflict with those for the *die*-frame. [12]

In the following, I will consider how the present findings could be captured by other accounts of framing effects (see Section 2), particularly the finding that the framing effect was absent when the numerals in the *survive*-version and the *die*-version were modified with an upward-entailing and downward-entailing comparative quantifier, respectively.

According to prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981), the source of framing effects lies in the different reference outcomes being induced by the alternative frames, such that an outcome that is described with the *survive*-frame is perceived as a gain whereas an outcome that is described with the *die*-frame is perceived as a loss. To account for the present findings within the framework of prospect theory, one might conjecture that the reference outcome, relative to which the outcome of a given option is evaluated, is not only determined by the *survive*-frame vs. *die*-frame, but crucially also by the presence of expressions such as comparative quantifiers. More specifically, the outcome description *more than 200 people will survive* might induce a reference outcome of 200 lives saved and the outcome description *fewer than 400 people will die* might induce a reference outcome of 400 lives lost, such that for both description variants, the outcome is perceived as a gain.

In fuzzy-trace theory (Reyna and Brainerd, 1991), framing effects are assumed to emerge from gist representations that do not contain the numerical information. To capture the present findings, one could speculate that the gist representations for the *survive*-version and the *die*-version of the sure option with modified numerals correspond to something like (19a) and

[11] The inferences for the *survive*-version and the *die*-version with modified numerals correspond to (ia) and (ib).

(i) a. 'more than $n + 1$ lives saved' is better than 'more than n lives saved'
b. 'fewer than n lives lost' is better than 'fewer than $n + 1$ lives lost'

[12] For bare numerals, the inferences for the *survive*-version and the *die*-version correspond to (ia) and (ib).

(i) a. '$n + 1$ lives saved' is better than 'n lives saved'
b. '$n + 1$ lives lost' is better than 'n lives lost'

(19b), respectively. However, to explain the present findings for modified numerals, it must additionally be assumed that the gist representations for the two versions do not give rise to (substantially) different valence evaluations.

(19) a. If Program A1 is adopted, more people will survive.
b. If Program B1 is adopted, fewer people will die.

The ambiguity-hypothesis (e.g. Kühberger and Tanner, 2010) claims that framing effects have their source in the incompleteness of the description of the sure option. It is assumed that the incomplete descriptions allow for wild guesses as to the fate of the unmentioned people with different outcomes for the two frames, such that the two options are not truth-conditionally equivalent. To capture the present findings within the framework of the ambiguity hypothesis, one either must assume that there are no wild guesses with modified numerals or that the guesses have similar outcomes for the *survive*-version and the *die*-version of the sure option.

The lower-bound-reading account of framing effects (Mandel, 2014) also attributes framing effects to a truth-conditional difference between the *survive*-version and the *die*-version of the sure options. More specifically, it is assumed that the numerals that indicate the amount of lives saved or lost receive a lower-bound reading. Apparently, this account can readily explain the present findings by the non-controversial assumption that the reading of numerals is constrained when they are modified by comparative quantifiers. However, to capture the present findings within the lower-bound-reading account, one has to make the additional assumption that the *survive*-version with an upward-entailing comparative quantifier and the *die*-version with a downward-entailing comparative quantifier are either equivalent or that that they differ less than the standard versions with bare numerals to which a lower-bound reading is assigned.

According to the information-leakage account (e.g. McKenzie and Nelson, 2003), framing effects arise because descriptions with different frames are not information equivalent. It is assumed that the *survive*-version and the *die*-version of the sure options convey different implicit information with regard to how their outcome deviates from a reference point. To cover the present findings, a core assumption of the account must be revised, i.e., the notion that solely the frame is decisive has to be abandoned. In order to explain the findings within the framework of the information-leakage account, one has to assume that expressions such as comparative quantifiers also bear on reference-point inferences. Regarding the case at hand, one might hypothesize that the outcome description *more than 200 people will survive* implicitly conveys a reference point of 200 surviving people and that the outcome description *fewer than 400 people will die* implicitly conveys a reference point of 400 dying people.

The argumentative-orientation account (Holleman and Pander Maat, 2009) also acts on the assumption that framing effects can be attributed to a difference in implicit information between distinct frames. It is assumed that different frames trigger different argumentative-orientation implicatures. To capture the present findings within the framework of the argumentative-orientation account calls for assumptions as to whether and how the presence of comparative quantifiers is pivotal in prompting argumentative-orientation implicatures.

In summary, the findings for modified numerals from the present study are consistent with predictions that directly follow from the alignment-assumption account. The findings can be captured by other accounts of framing effects, albeit only by making additional, post-hoc assumptions. An exception in this regard is the lower-bound-reading account. However, as was mentioned in Section 2.2, studies on the role of numeral interpretations for framing effects yielded mixed results. It remains an empirical task to reconcile the equivocal findings and to further explore the validity of the lower-bound-reading account.

5. Conclusion

Framing effects have been demonstrated in numerous experimental studies. Yet, their source is still a matter of controversy. There is a wide range of different accounts of framing effects. The accounts are not mutually incompatible, and it is likely that multiple factors may be involved. The very basis of framing effects, however, are differences in linguistic input. Hence, pushing the advancement of semantic-pragmatic explanations is highly desirable. The alignment-assumption account proposed by Geurts (2009, 2013) offers such an explanation. It radically differs from other accounts and opens up a new perspective on framing effects by explaining them in matters of a counterfactual systematicity of judgments and choices. The findings of the present experiments provide first empirical support for the alignment-assumption account by confirming two directly derivable predictions. The predictions, which pertained to effects of upward- vs. downward entailing comparative quantifiers, do not follow from other accounts of framing effects, which are challenged by the present findings.

Obviously, the scope of the present study is limited. It remains to be investigated whether people actually act from the alignment assumption[13] and whether or not the assumption is an internalized heuristic. Such kind of investigations are specifically relevant as to the validity of the alignment-assumption account. However, their findings might be of wider relevance, i.e., to research on the semantics of evaluative predicates and to research on entailment reasoning.

References

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25*(2), 93–139.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston and S. Uchida (Eds.), *Relevance theory: Applications and implications*, pp. 179–236. Amsterdam: John Benjamins.

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.

Druckman, J. N. (2001). Evaluating framing effects. *Journal of Economic Psychology, 22*(1), 91–101.

[13] The initial results of an own ongoing pilot study are promising in this regard. Participants overwhelmingly agreed with statements that corresponded to alignment-assumption inferences from evaluative assertions on fictitious issues for which they had no prior knowledge (e.g. *It is good that 36 of the tested shampoos contained Burarlin*), and they overwhelmingly disagreed with statements that were in conflict with the inferences.

Geurts, B. (2006). Take "five": The meaning and use of a number word. In S. Vogeleer and L. Tasmowski (Eds.), *Non-definiteness and plurality*, pp. 311–329. Philadelphia, PA: Benjamins.

Geurts, B. (2009). Goodness. In M. Aloni, H. Bastiaanse, T. de Jager, P. van Ormondt, and K. Schulz (Eds.), *Preproceedings of the 17th Amsterdam Colloquium*, pp. 277–285. Amsterdam: ILLC/Department of Philosophy, University of Amsterdam.

Geurts, B. (2013). Alternatives in framing and decision making. *Mind & Language, 28*(1), 1–19.

Holleman, B. C. and H. L. W. Pander Maat, (2009). The pragmatics of profiling: Framing effects in text interpretation and text production. *Journal of Pragmatics, 41*(11), 2204–2221.

Horn, L. (1992). The said and the unsaid. In C. Barker and D. Dowty, *Proceedings of Semantics and Linguistic Theory 2*, pp. 163–192. Columbus, OH: OSU Working Papers in Linguistics.

Kennedy, C. (2015). A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics, 8*, 1–44.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Krifka, M. (2009). Approximate interpretations of number words: a case for strategic communication. In E. Hinrichs and J. Nerbonne (Eds.), *Theory and Evidence in Semantics,* pp. 109–132). Stanford, CA: CSLI Publications.

Kühberger, A. (1995). The framing of decisions: A new look at old problems. *Organizational Behavior and Human Decision Processes, 62*(2), 230–240.

Kühberger, A. and P. Gradl (2013). Choice, rating, and ranking: Framing effects with different response modes. *Journal of Behavioral Decision Making, 26*(2), 109–117.

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.

Levin, I. P., S. L. Schneider, and G. J. Gaeth (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76*(2), 149–88.

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.

Mandel, D. R. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General, 143*(3), 1185–1198.

McKenzie, C. R. M. and J. D. Nelson (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin and Review, 10*(3), 596–602.

Mercier, H. and D. Sperber (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57–111.

Peters, E. and I. P. Levin (2008). Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options. *Judgment and Decision Making, 3*(6), 435–448.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making, 7*(3), 332–359.

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.

Setton, R. A., E. A. Wilhelms, R. B. Weldon, C. F. Chick, and V. F. Reyna (2014). An overview of judgment and decision making research through the lens of fuzzy trace theory. *Advances in Psychological Science, 22*(12), 1837–1854.

Sher, S. and C. R. M. McKenzie (2006). Information leakage from logically equivalent frames. *Cognition, 101*(3), 467–494.

Tversky, A. and D. Kahneman (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Werner, K. M. and H. Zank (2018). A revealed reference point for prospect theory. *Economic Theory*, https://doi.org/10.1007/s00199-017-1096-2.