

# Framing in context: Disabling conditions and alternative causes in health communication

**Peter J. Collins** (p.j.collins@greenwich.ac.uk)

School of Human Sciences, University of Greenwich  
Dreadnought Building, Park Row, London SE10 9NN, United Kingdom

**Karolina Krzyzanowska** (k.h.krzyzanowska@uva.nl)

Philosophy Department & Institute for Logic, Language and Computation, University of Amsterdam  
Science Park 107, 1098 XG Amsterdam, The Netherlands

**Ulrike Hahn** (u.hahn@bbk.ac.uk)

Department of Psychological Sciences, Birkbeck, University of London  
Malet Street, London WC1E 7HX, United Kingdom

## Abstract

Should a health campaign emphasise the potential gains from compliance (e.g., “If you quit smoking, you’ll reduce your risk of lung cancer”) or the potential losses from non-compliance (e.g., “If you don’t quit smoking, you won’t reduce your risk of lung cancer”)? A large literature on so-called goal framing, or message framing, assumes that such messages are equivalent, but their persuasiveness may vary, for instance, depending on the perceived risk associated with the recommended behaviour. As no existing hypothesis received conclusive empirical support, we propose a novel theoretical approach. We argue that goal frames must be analysed as arguments interpreted in context. We report an experiment showing the effect of the participants’ background beliefs about disabling conditions and alternative causes on the persuasiveness of positive and negative frames recommending detection behaviour.

**Keywords:** goal framing; conditionals; arguments from consequences; acceptability; attitudes; health communication

## Introduction

Should a health campaign emphasise the potential gains from compliance (e.g., “If you quit smoking, you’ll reduce your risk of lung cancer”) or the potential losses from non-compliance (e.g., “If you don’t quit smoking, you won’t reduce your risk of lung cancer”)? The two presentations of what is assumed to be the same information are called *frames* or, more specifically, *goal frames*.<sup>1</sup> Since Tversky & Kahneman’s seminal 1981 paper, whenever equivalent frames trigger different responses of their recipients, we talk about *framing effects*. More recently, it has been suggested that for a framing effect to arise, messages should not only be logically equivalent but also, preferably, they should convey the same information content (see, e.g., Sher & McKenzie 2006; Corner & Hahn 2010).

Positively and negatively framed messages used in health communication typically take the form of indicative conditionals, “If  $p$ ,  $q$ ” and “If not  $p$ , not  $q$ ,” where  $p$  and  $q$  are typ-

ically simple propositions.<sup>2</sup> More specifically, goal frames can be schematically represented as:

- (1) If you A, you will C.
- (2) If you do not A, you will not C.

Here, A stands for the advocated action or behaviour, like quitting smoking, exercising, or getting tested for a virus, while C stands for the health benefits that are supposed to follow. A large literature on goal framing, or message framing, holds that such messages are “objectively equivalent” (Van ’t Riet et al. 2016, p. 456) or “conveying essentially identical information” (Gallagher & Updegraff 2012, p. 101). At the same time, differently framed messages are expected to have different persuasive effects on their recipients.

Importantly, however, despite the prevalence of the equivalence assumption in the goal-framing literature, on no semantic or psychological account of conditionals are (1) and (2) logically equivalent (e.g., Bennett 2003; Egré & Rott 2021; Evans & Over 2004). Following the standard definition, two propositions are logically equivalent if they are true in exactly the same situations (and false in exactly the same situations). Yet the fact that “If you swim regularly, you are physically active” is true does not guarantee that so is “If you don’t swim regularly, you aren’t physically active”; an individual who is regularly running, cycling, or lifting weights is physically active too. It has been observed that some contexts seem to invite an inference from a conditional to its inverse. For instance, “If you do the dishes, I’ll buy you ice cream” said by a parent to a child is likely to be interpreted as conveying “If you don’t do the dishes, I won’t buy you ice cream.” While there has been some debate about the exact mechanism underlying this inference, its pragmatic nature is uncontroversial (see, e.g., Geis & Zwicky 1971; Horn 2000; van der Auwera 1997; Noveck 2018, ch. 8).

<sup>2</sup>While messages such as “Exercising regularly can help you lose weight” do not share the surface form with a standard indicative conditional, they can be seen as expressing conditional thoughts nonetheless; for a helpful discussion see, e.g., Elder & Jaszczolt 2016.

<sup>1</sup>The terms “frame” and “framing” in different contexts can refer to fundamentally different phenomena and cannot be captured by a single definition (Bermúdez 2020). In this paper, we focus exclusively on *goal frames*, which is a distinct category within the class of equivalence frames (see, e.g., Levin et al. 1998). For a broader overview of different types of frames see, e.g., Flusberg et al. (2024).

Since gain- and loss-framed messages are not even equivalent, they cannot be expected to convey exactly the same information content. Consequently, they are not strictly speaking frames at all<sup>3</sup> (cf. Corner & Hahn 2010). Both (1) and (2) can nonetheless be seen as arguments for A, thus the question of whether a positively or a negatively framed conditional is more effective in promoting healthy behaviours remains valid and important.

It has been hypothesised, for instance, that negative frames, i.e., those that emphasise potential losses, should be more persuasive due to people's negativity bias (Meyerowitz & Chaiken 1987; O'Keefe & Jensen 2006), that is, the tendency to pay more attention or to attach greater weight to negative information (Rozin & Royzman 2001). Persuasiveness of negative frames could also be underpinned by loss aversion (Kahneman & Tversky 1979). However, as revealed in a meta analysis by O'Keefe & Jensen (2006), there is little empirical support for the negativity hypothesis, as loss frames are not generally more persuasive than gain frames.

By contrast, Rothman & Salovey (1997), inspired by Tversky & Kahneman (1981), suggested that persuasiveness of a given type of frame should depend on the perceived short-term risk associated with a given behaviour. In the context of health-related messages, risk is understood as the chances of a negative outcome or, in other words, "danger." Understood this way, prevention behaviours such as exercise or healthy eating are low-risk, while detection behaviours, such as cancer screening or HIV testing, are high-risk in the short-term: they bring about a danger of finding out that one has a disease or carries a potentially deadly virus. Following Tversky and Kahneman's discovery that "choices involving gains are often risk averse and choices involving losses are often risk taking" (Tversky & Kahneman 1981, p. 458), Rothman & Salovey (1997) proposed that positive messages, that is gain frames, should be more persuasive for the prevention behaviours (analysed as low-risk) while negative messages, that is loss frames, should be more effective for the detection behaviours (analysed as high-risk).<sup>4</sup>

Empirical findings on risk matching are, however, generally inconclusive. A meta-analysis by O'Keefe & Jensen (2007) demonstrated that gain-frames are not reliably more persuasive than loss-frames for most prevention behaviour, except in the context of dental hygiene messaging. Similarly, a meta-analysis by O'Keefe & Jensen (2009) revealed that loss frames are not reliably more persuasive than gain frames for most detection behaviour, with an exception of breast cancer detection messaging. Gallagher & Updegraff (2012) found a small but significant effect of gain framing on prevention behaviour, but not for other measures such as

<sup>3</sup>For consistency with the existing literature, we continue to use the term "frame" to refer to such messages.

<sup>4</sup>Van 't Riet et al. (2016), following O'Keefe & Jensen (2006), argued that this simple version of risk-matching hypothesis is a misapplication of Prospect Theory (Kahneman & Tversky 1979), in the context of which risk is understood in probabilistic terms. Nevertheless, we include the simple risk-matching hypothesis in our discussion as it plays a prominent role in the goal-framing literature.

attitudes or intentions to behave. This heterogeneity of the effects calls for "new strategies that might enable [the investigators] to disentangle the signal from the noise" (Rothman et al. 2020, p. 9). To that end, we propose to take a step back and analyse goal frames as linguistic expressions that are asserted and interpreted in a context. More specifically, we take goal frames to be arguments from consequences which are often phrased as indicative conditionals. This new theoretical approach allows us to draw from the large empirical literature devoted to conditionals and conditional reasoning.

## Conditional reasoning

Decades of psychological and linguistic research<sup>5</sup> demonstrated that participants' interpretation of conditionals and inferences they draw from conditional premisses are susceptible to various contextual factors such as speakers' expertise (Stevenson & Over 2001; Collins et al. 2020), but also the participants' own background beliefs. For instance, in reasoning tasks in which the conditional premiss expresses a causal relation, participants' judgments are affected by the presence of *disabling conditions* and *alternative causes* (Cummins et al. 1991; Cummins 1995). Disabling conditions (or *disablers* for short) are circumstances that prevent the effect from occurring. For instance, while eating candy often is a cause of cavities, rigorous dental hygiene may prevent it from coming about. Similarly, eating candy is not the only possible cause of cavities; frequent consumption of sugary drinks or lack of dental hygiene are alternative causes (or *alternatives* for short) that may result in the same effect.

Since conditionals used in health communication are meant to convey that the action recommended in the antecedent would lead to the health benefits promised in the consequent, the way the target audience responds to these messages may also be affected by their own beliefs about the presence of disablers and alternatives. For instance, when a health campaign advocates quitting smoking by emphasising gains that are supposed to follow, the audience's background beliefs in the presence of disablers may undermine the message's persuasiveness. Someone could doubt that quitting smoking would result in an improvement of their health if they live in a highly polluted city. A strong belief that smog exposure will nullify any possible gains renders the conditional "If you quit smoking, you reduce your risk of lung cancer" highly unacceptable (or false), and thus a poor argument for quitting smoking. Similarly, disabling conditions may prevent a test from delivering reliable results, making thus a gain-framed message unacceptable and thus less persuasive.

The presence of alternative causes should not affect the acceptability of the conditional itself: (1) is not false if there are other ways to bring about C. The negative frame (2), however, communicates that A is necessary to achieve C, which would be undermined by the presence of alternatives. Loss-framed messages such as "If you don't quit smoking, you

<sup>5</sup>For an introduction to the field, see Evans & Over 2004; and for an overview of more recent debates, see, e.g., Kaufmann et al. 2023.

won't reduce your risk of lung cancer" emphasise that giving up smoking is necessary to achieve the health benefits mentioned in the consequent. The presence of alternative ways to achieve the same goal, or rather, the audience's beliefs that alternatives are available, will undermine the perceived necessity of the connection between the advocated action and its effects, and, consequently, the persuasiveness of the message.

## The hypotheses

The experiment reported below belongs to a broader project on message framing, involving studies on different types of behaviours and several dependent variables. Here we report a pre-registered study (<https://osf.io/ytcf2/>) on the effect of disablers and alternatives on the persuasiveness of positive and negative frames recommending detection behaviour and their acceptability in a context.

More specifically, from research on causal reasoning and conditionals (Cummins et al. 1991; Cummins 1995) we derive the *background-beliefs hypothesis* which we will contrast with the simple risk-matching hypothesis (Rothman & Salovey 1997) and negativity bias/loss aversion hypothesis (Meyerowitz & Chaiken 1987). While meta analyses demonstrated these hypotheses lack empirical support, the risk framing hypothesis is still highly influential (see, e.g., Nan et al. 2018). As the frames in our experimental materials advocate detection behaviour, the simple risk-matching hypothesis predicts negative frames to be more persuasive, resulting in higher ratings for both Acceptability and Attitude. The same prediction can be derived from the negativity bias/loss aversion hypothesis, which we include here for completeness.

By contrast, according to the background-belief hypothesis, framing should interact with the number of disablers and with the number of alternatives. On the one hand, we expect the number of disablers to have a more pronounced effect with positive frames. On the other hand, we expect the number of alternatives to have a more pronounced effect with the negative frames. Moreover, we expect that both fewer disablers and fewer alternatives will lead to more positive attitudes towards the advocated action.

## The Experiment

We developed a medical scenario about a fictional disease called "Colin Syndrome" (modelled on Cushing Syndrome). As typical in the framing literature, the vignette emphasises the importance of early detection to allow the best forms of treatment. In the experiment, we ask the participants to imagine themselves in the given scenario to make sure that they consider the messaging as relevant for their situation. The experimental materials advertise a detection method that should be familiar to the participants: collecting saliva with a swab test. The participants are also informed that the swabs are uncomfortable to take but not painful.

The study follows  $2 \times 2 \times 2$  design, with Framing (Positive/Negative), Disablers (None/Many), and Alternatives

(None/Many) manipulated between participants. The vignette provides an argument for taking the swab test by emphasising either the desirable consequences of compliance (positive frame) or the undesirable consequences of non-compliance (negative frame):

**Positive frame:** If you take a swab test, you will know early in your illness whether you have Colin Syndrome.

**Negative frame:** If you don't take a swab test, you will not know early in your illness whether you have Colin Syndrome.

To manipulate the participants' background beliefs about the number of disablers and the number of alternatives, we provide them with the following information:

**No (many) disablers:** Taking a swab test guarantees (does not guarantee) reliable and conclusive results. That's because the swabs are easy (hard) to collect, the testing is unaffected (affected) by personal characteristics, and processing of the swabs is straightforward (difficult).

**No (many) alternatives:** No (several) other effective ways of testing for Colin Syndrome are available. For example, it is not possible (it is possible) to test using blood, urine, or stool samples.

We explore the effects of these variables on the perceived acceptability of the framed statement and on attitudes towards the advocated behaviour as measured on 7-point semantic differentials taken from Buda & Zhang (2000) and, as manipulation checks, on measures of perceived sufficiency and necessity of the behaviour for the effect.

## Method

**Participants.** The participants were recruited via the Prolific platform (<https://www.prolific.com/>) and received remuneration for completing the survey. Analyses<sup>6</sup> are conducted on those participants ( $n = 716$ , mean age 42.92,  $sd = 13.91$ ) who provided answers to all questions and who passed the attention test. 310 participants identified as Male/Man, 395 as Female/Woman, 3 as Non-binary, 4 preferred to self-describe and 4 preferred not to say. Note that owing to an error with the Prolific settings, a number of participants participated who did not have English as a first language and a number participated multiple times. These participants were excluded.

**Materials and procedure.** After giving informed consent and providing demographics information, participants were randomly assigned to one of 8 conditions. After reading the instructions explaining the nature of the experiment, participants in all conditions were asked to imagine themselves in the following medical scenario:

<sup>6</sup>All analyses other than those of the interaction of disablers and alternatives and their follow ups are pre-registered.

| Effect       | <i>df</i> | <i>MSE</i> | <i>F</i>  | <i>ges</i> | <i>p</i> |
|--------------|-----------|------------|-----------|------------|----------|
| Fr           | 1,708     | 2.60       | 2.41      | .003       | .12      |
| Dis          | 1,708     | 2.60       | 156.02*** | .18        | < .001   |
| Alt          | 1,708     | 2.60       | 104.07*** | .13        | < .001   |
| Fr: Dis      | 1,708     | 2.60       | 59.46***  | .08        | < .001   |
| Fr: Alt      | 1,708     | 2.60       | 64.53***  | .08        | < .001   |
| Dis: Alt     | 1,708     | 2.60       | 0.28      | < .001     | .60      |
| Fr: Dis: Alt | 1,708     | 2.60       | 0.33      | < .001     | .57      |

Table 1: Three-way ANOVA on Acceptability

You have started experiencing some minor symptoms which are consistent with a condition called Colin Syndrome. This is a condition which leads to bruising, weakness in the limbs, and mood swings. It also leads to high blood pressure, which is serious if untreated. The condition is treatable, but it is important to find out early in your illness whether you have Colin Syndrome to allow the best forms of treatment where necessary.

Next, participants received information about the possibility of testing for the disease by collecting saliva with a swab test, followed by additional information pertaining to the number of disablers and the number of alternatives. Next, participants were asked to read the statement providing an argument for the swab test (i.e., a positive or a negative frame), which was followed by the test questions measuring the following variables:

**Perceived acceptability** (“How acceptable is this statement in this context?”) was measured on a 7-point scale from 1 (*Highly Unacceptable*) to 7 (*Highly Acceptable*).

**Attitudes** towards the swab test were measured on nine 7-point semantic differential scales arranged in three sets:

1. “How attractive do you find the swab test as an option?” (*bad / good, not nice / nice, and unlikeable / likeable.*)
2. “Would you take the swab test for Colin Syndrome?” (*unlikely / likely, improbable / probable, impossible / possible.*)
3. “How confident are you that the swab test would work?” (*not confident / confident, not certain / certain, feel unsure / feel sure.*)

Additionally, participants were presented with an attention check question and two manipulation check questions.

## Results: Acceptability

This section reports results in the order of the hypotheses. Table 1 shows the results of the three-way ANOVA on the acceptability data. As the three-way interaction was not significant, we precede with the lower-level effects.

**Main Effect of Framing** If the loss aversion hypothesis, negativity bias or a simple version of the risk-matching hypotheses hold, then negative frames should be more acceptable than positive frames. In fact, the main effect of Framing

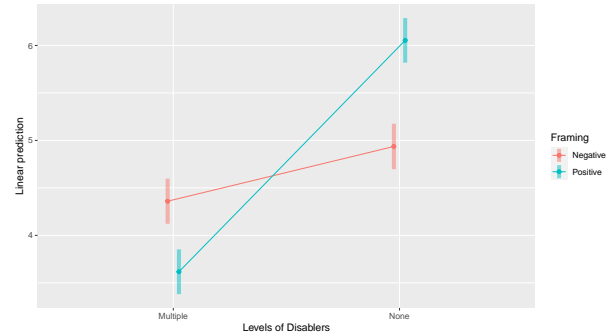


Figure 1: Estimated marginal means showing the interaction of Framing and Disablers on Acceptability. The error bars represent 95% Confidence Intervals.

was not significant, with the estimate marginal means showing very similar ratings for positive frames ( $M = 4.84$ ,  $SE = .08$ , 95% CI [4.67, 5.00]) and negative frames ( $M = 4.65$ ,  $SE = .09$ , 95% CI [4.48, 4.82]). Framing, additionally, entered into significant two-way interactions.

**Interaction of Framing and Disablers** As predicted under the background-beliefs hypothesis, there was a significant interaction of framing and disablers, shown in Figure 1, which indicates a cross-over interaction. Simple effects analyses were treated as a family of four tests, with the significance level set to  $p < .013$ . While having no rather than multiple disablers consistently improved acceptability, the effect was rather larger for positive frames ( $M_{Difference} = 2.44$ ,  $SE = .17$ ,  $t(708) = 14.38$ ,  $p < .001$ ) than for negative frames ( $M_{Difference} = 0.58$ ,  $SE = .17$ ,  $t(708) = 3.36$ ,  $p < .001$ ). When there were multiple disablers, negative frames were rated significantly more acceptable than positive frames ( $M_{Difference} = .74$ ,  $SE = .17$ ,  $t(708) = 4.36$ ,  $p < .001$ ). When there were no disablers, positive frames were rated significantly more acceptable than negative frames ( $M_{Difference} = 1.12$ ,  $SE = .17$ ,  $t(708) = 6.54$ ,  $p < .001$ ). These effects support the background-beliefs hypothesis.

**Interaction of Framing and Alternatives** As predicted under the background-beliefs hypothesis, there was a significant interaction of Framing and Alternatives, shown in Figure 2 which indicates a crossover interaction.

Simple effects analyses were again treated as a family of four tests, with the significance level set to  $p < .013$ . While having no rather than multiple alternatives appeared to improve ratings of acceptability, this effect was only significant for negative frames ( $M_{Difference} = 2.20$ ,  $SE = .17$ ,  $t(708) = 12.81$ ,  $p < .001$ ) with only a small difference for positive frames ( $M_{Difference} = .26$ ,  $SE = .17$ ,  $t(708) = 1.54$ ,  $p = .12$ ). When there were multiple alternatives, positive frames were rated significantly more acceptable than negative frames ( $M_{Difference} = 1.16$ ,  $SE = .17$ ,  $t(708) = 6.79$ ,  $p < .001$ ). When there were no alternatives, negative frames were rated significantly more acceptable than positive frames

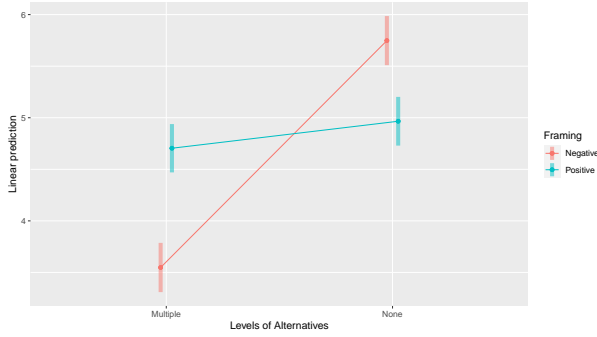


Figure 2: Estimated marginal means showing the interaction of Framing and Alternatives on Acceptability. The error bars represent 95% Confidence Intervals.

( $M_{Difference} = .78, SE = .17, t(708) = 4.58, p < .002$ ).

**Main Effects of Disablers and Alternatives** These effects should be interpreted in light of the significant interactions into which the variables enter. As the estimated marginal means show, the effect of Disablers was such that, on average, claims were rated as more acceptable when there were no disablers ( $M = 5.49, SE = .09, 95\% CI [5.33, 5.55]$ ) than when there were multiple disablers ( $M = 3.99, SE = .09, 95\% CI [3.82, 4.16]$ ). The effect of Alternatives was such that, on average, claims were rated as more acceptable when there were no alternatives ( $M = 5.36, SE = .09, 95\% CI [5.19, 5.52]$ ) than when there were multiple alternatives ( $M = 4.13, SE = .09, 95\% CI [3.96, 4.29]$ ). As detailed above, however, since the effect of Alternatives was not significant for positive frames, there is somewhat qualified support for our background-beliefs hypothesis.

### Results: Attitude

Table 2 shows the results of the three-way ANOVA conducted on the attitude data. The dependent variable is the average of the semantic differential scales. As the three-way interaction was not significant, we proceed with the lower-level effects.

**Main Effect of Framing** If the loss aversion hypothesis or a simple version of the risk-matching hypotheses hold, then negative frames should result in more favourable attitudes than positive frames. The main effect of Framing was not sig-

| Effect       | <i>df</i> | <i>MSE</i> | <i>F</i>  | <i>ges</i> | <i>p</i> |
|--------------|-----------|------------|-----------|------------|----------|
| Fr           | 1, 708    | 1.13       | 3.33      | .005       | .07      |
| Dis          | 1, 708    | 1.13       | 322.40*** | .31        | < .001   |
| Alt          | 1, 708    | 1.13       | 19.64***  | .03        | < .001   |
| Fr: Dis      | 1, 708    | 1.13       | 5.33*     | .007       | .02      |
| Fr: Alt      | 1, 708    | 1.13       | .32       | < .001     | .57      |
| Dis: Alt     | 1, 708    | 1.13       | 31.86***  | .04        | < .001   |
| Fr: Dis: Alt | 1, 708    | 1.13       | 0.29      | < .001     | .59      |

Table 2: Three-way ANOVA on Attitude (Detection).

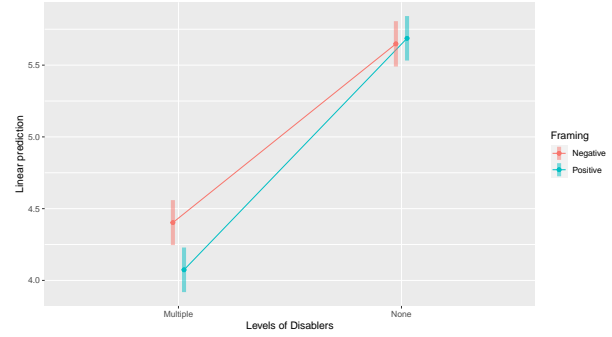


Figure 3: Estimated marginal means showing the interaction of Framing and Disablers on Attitude. The error bars represent 95% Confidence Intervals.

nificant, but the estimated marginal means were in line with that prediction, with negative frames ( $M = 5.03, SE = .06, 95\% CI [4.91, 5.14]$ ) producing slightly more favourable attitudes than positive frames ( $M = 4.88, SE = .06, 95\% CI [4.77, 4.99]$ ). The main effect of framing was, however, qualified by a significant two-way interaction with Disablers.

**Interaction of Framing and Disablers** As predicted under the background-beliefs hypothesis, there was a significant interaction of Framing and Disablers, shown in Figure 3.

Simple effects analyses were treated as a family of four tests, with the significance level set to  $p < .013$ . While having no rather than multiple disablers consistently improved attitude, the effect was somewhat larger for positive frames ( $M_{Difference} = 1.61, SE = .11, t(708) = 14.43, p < .001$ ) than for negative frames ( $M_{Difference} = 1.25, SE = .11, t(708) = 10.99, p < .001$ ). When there were multiple disablers, negative frames produced significantly higher attitude than positive frames ( $M_{Difference} = .33, SE = .11, t(708) = 2.93, p = .004$ ). When there were no disablers, positive frames produced attitudes slightly but non-significantly higher than negative frames ( $M_{Difference} = .04, SE = .11, t(708) = 0.34, p = .73$ ). These data are in line with the background-beliefs hypothesis.

**Lack of Interaction of Framing and Alternatives; Main effect of Alternatives** Contrary to predictions under the background-beliefs hypothesis, there was no significant interaction between Framing and Alternatives. Indeed, there was an unqualified main effect of Alternatives (see below).

**Main Effects of Disablers and Alternatives** The main effect of Disablers should be interpreted in light of the significant interaction into which the variable enters. As the estimated marginal means show, the effect of Disablers was such that, on average, attitude scores were higher when there were no disablers ( $M = 5.67, SE = .06, 95\% CI [5.56, 5.78]$ ) than when there were multiple disablers ( $M = 4.24, SE = .06, 95\% CI [4.13, 4.35]$ ). As the estimated marginals also show, the effect of Alternatives was such that, on average, attitude scores

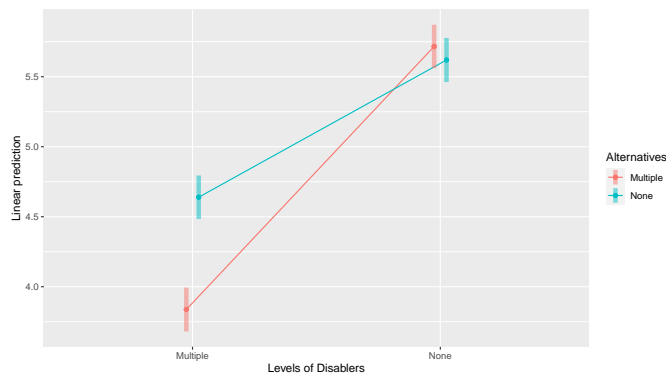


Figure 4: Estimated marginal means showing the interaction of Disablers and Alternatives on Attitude. The error bars represent 95% Confidence Intervals.

were higher when there were no alternatives ( $M = 5.13$ ,  $SE = .06$ , 95% CI [5.02, 5.24]) than when there were multiple alternatives ( $M = 4.78$ ,  $SE = .06$ , 95% CI [4.67, 4.89]).

**Exploratory Analysis: Interaction of Disablers and Alternatives** There was an unpredicted interaction of Disablers and Alternatives. This analysis is treated as exploratory; hence, no judgements are made regarding significance and no corrections are applied.

For both levels of the Alternatives variable, having no disablers led to higher attitude scores than having multiple disablers. However, when there were multiple alternatives, the effect of Disablers appeared slightly larger ( $M_{Difference} = 1.88$ ,  $SE = .11$ ,  $t(708) = 16.71$ ,  $p < .001$ ) than when there were no alternatives ( $M_{Difference} = .98$ ,  $SE = .11$ ,  $t(708) = 8.69$ ,  $p < .001$ ). The effect of Alternatives was less consistent. When there were multiple disablers, there was a clear advantage for no over multiple alternatives ( $M_{Difference} = .80$ ,  $SE = .11$ ,  $t(708) = 7.14$ ,  $p < .001$ ); when there were no disablers, there was little difference between Alternatives conditions, with multiple alternatives leading to slightly but non-significantly higher attitude scores than no alternatives. ( $M_{Difference} = .10$ ,  $SE = .11$ ,  $t(708) = .86$ ,  $p = .39$ ). These findings may suggest that, when a context is weaker with respect to one factor, the effect of the other factor is emphasised.

**Summary** Contrary to the predictions of the loss aversion hypothesis and the simple risk-matching hypothesis, negative frames were not generally more persuasive: positive and negative frames received very similar Acceptability ratings. While negative frames seemed to result in slightly more favourable attitudes towards the advocated action, that effect was not significant. The Acceptability data broadly supported the background-beliefs hypothesis: framing interacted with disablers and with alternatives. Participants found positive frames more acceptable than negative frames when there were no disablers while negative frames were more acceptable than positive frames when there no alternatives. For

these data, though, the effect of Alternatives was less consistent than expected, not significantly improving acceptability scores with positive frames. The data were less supportive of our hypotheses for Attitude. As predicted, we did observe that Disablers had a more pronounced effect on Attitude with positive frames. By contrast, there was no significant interaction between Framing and Attitude. That disablers consistently and significantly improved attitude scores for both frames (attested in the simple effects) and that there was a main effect of Alternatives are consistent with our predictions.

## Discussion

This paper set out to explore an alternative theoretical approach to goal framing, grounded in theories of conditionals and causal reasoning. Conventional approaches such as loss aversion, negativity bias, and the risk-matching hypothesis have met with inconsistent results. These approaches assume that the frames are logically equivalent. Our alternative approach rejects this assumption and, instead, emphasizes the importance of background beliefs about disabling conditions and alternative causes to how a conditional is interpreted and received.

Our data are broadly consistent with this alternative approach, suggesting that disabling conditions matter more to the acceptability of positive frames and alternatives matter more to the acceptability of negative frames. Importantly, different acceptability ratings for positive and negative frames indicate that people do not perceive the frames as “conveying essentially identical information” (Gallagher & Updegraff 2012) or being “objectively equivalent” (Van ’t Riet et al. 2016). Quite the contrary; the discrepancy between the acceptability ratings for negative and positive frames in all conditions suggests that people are sensitive to semantic differences between the two conditionals. A similar, but less strongly supportive (with one missing interaction), pattern is seen with the attitude data.

These data require further replication with different dependent variables and the background-beliefs hypothesis requires extending to different – for example, prevention – behaviours. However, we suggest that an approach grounded in theories of conditionals and causal reasoning is a promising alternative to existing theoretical approaches and may offer a valuable way to tease apart the signal and the noise in data on goal framing.

## References

- Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.
- Bermúdez, J. L. (2020). *Frame it again: New tools for rational decision-making*. Cambridge University Press.
- Buda, R., & Zhang, Y. (2000). Consumer product evaluation: the interactive effect of message framing, presentation order, and source credibility. *Journal of Product & Brand Management*, 9(4), 229-242.

- Collins, P. J., Krzyżanowska, K., Hartmann, S., Wheeler, G., & Hahn, U. (2020). Conditionals and testimony. *Cognitive Psychology*, *122*, 101329.
- Corner, A., & Hahn, U. (2010). Message framing, normative advocacy and persuasive success. *Argumentation*, *24*(2), 153-163.
- Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, *23*(5), 646-658.
- Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, *19*(3), 274-282.
- Egré, P., & Rott, H. (2021). The Logic of Conditionals. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2021 ed.). Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/win2021/entries/logic-conditionals/>
- Elder, C.-H., & Jaszczolt, K. M. (2016). Towards a pragmatic category of conditionals. *Journal of Pragmatics*(98), 36-53.
- Evans, J. S. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Flusberg, S. J., Holmes, K. J., Thibodeau, P. H., Nabi, R. L., & Matlock, T. (2024). The psychology of framing: How everyday language shapes the way we think, feel, and act. *Psychological Science in the Public Interest*, *25*(3), 105-161.
- Gallagher, K. M., & Updegraff, J. A. (2012). Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine*, *43*(1), 101-116.
- Geis, M. L., & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry*, *2*(4), 561-566.
- Horn, L. R. (2000). From *if* to *iff*: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics*, *32*, 289-326.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291.
- Kaufmann, S., Over, D. E., & Sharma, G. (Eds.). (2023). *Conditionals: Logic, linguistics and psychology*. Palgrave Macmillan.
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149-188.
- Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*(3), 500-510.
- Nan, X., Daily, K., & Qin, Y. (2018). Relative persuasiveness of gain- vs. loss-framed messages: A review of theoretical perspectives and developing an integrative framework. *Review of Communication*, *18*(4), 370-390.
- Noveck, I. (2018). *Experimental pragmatics*. Cambridge: Cambridge University Press.
- O'Keefe, D. J., & Jensen, J. D. (2006). The advantages of compliance or the disadvantages of noncompliance? a meta-analytic review of the relative persuasive effectiveness of gain-framed and loss-framed messages. *Communication Yearbook*, *30*(1), 1-43.
- O'Keefe, D. J., & Jensen, J. D. (2007). The relative persuasiveness of gain-framed loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review. *Journal of Health Communication*, *12*(7), 623-644.
- O'Keefe, D. J., & Jensen, J. D. (2009). The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease detection behaviors: A meta-analytic review. *Journal of Communication*, *59*(2), 296-316.
- Rothman, A. J., Desmarais, K.-J., & Lenne, R. L. (2020). Chapter two - moving from research on message framing to principles of message matching: The use of gain- and loss-framed messages to promote healthy behavior. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 7, p. 43-73). Elsevier.
- Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, *121*(1), 3-19.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296-320.
- Sher, S., & McKenzie, C. R. (2006). Information leakage from logically equivalent frames. *Cognition*, *101*(3), 467-494.
- Stevenson, R. J., & Over, D. E. (2001). Reasoning from uncertain premises: Effects of expertise and conversational context. *Thinking & Reasoning*, *7*(4), 367-390.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453-458.
- van der Auwera, J. (1997). Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics*, *27*(3), 261-274.
- Van 't Riet, J., Cox, A. D., Cox, D., Zimet, G. D., De Bruijn, G.-J., Van den Putte, B., ... Ruiters, R. A. C. (2016). Does perceived risk influence the effects of message framing? revisiting the link between prospect theory and message framing. *Health Psychology Review*, *10*(4), 447-459.