

Framing Affects Post-Decision Preferences Through Self-Preference Inferences
(and Probably not Dissonance)

*****Forthcoming in Journal of Experimental Psychology: General*****

Adelle X. Yang

Assistant Professor, National University of Singapore

adelle.yang@gmail.com

Jasper Teow

Lecturer, The Business School, RMIT University Vietnam

jasper.teow@rmit.edu.vn

Author Note

Correspondence should be addressed to Adelle X. Yang (ORCID: <https://orcid.org/0000-0001-5817-7050>), adelle.yang@gmail.com. The authors wish to thank Tom Meyvis, Leif Nelson, Jane Risen, Klaus Werternbroch, Oleg Urminsky, Clayton Critcher, and Reid Hastie for providing valuable comments on this research. We also acknowledge the helpful feedback from the editors, anonymous reviewers, attendees at the 2020 SJDM conference and the Singapore Island Labs. All data, original materials, and analysis code are shared on Open Science Framework at <https://osf.io/qhcdy/>. This research adheres to the American Psychological Association's ethical standards concerning participant treatment. This paper has not been shared on any academic social networks. The authors declare no conflict of interest.

Abstract (248/250)

Psychologists have long been intrigued by decision-induced changes in preferences, where making a decision strengthens one's relative preference between more and less preferred options. This phenomenon has been explained through two prominent theories: a dissonance account, which suggests that it results from the decision-maker's attempt to minimize an unpleasant emotional-motivational state of "dissonance," and an inference account, which posits that it reflects a process of inferring and updating one's "true" preferences. In the current research, we investigate whether, how, and why framing a decision as a choice or a rejection influences decision-induced preference modulation. Across thirteen pre-registered experiments including seven ($N = 6,248$ participants from North America and Asia) reported in the main text, we find that reject-framed decisions between attractive options induce greater post-decision preference modulation (i.e., a larger preference gap between options) than choose-framed decisions, all else equal. Supporting the inference account, the effect is moderated by attribute similarity and choice-set valence while mediated consistently by perceived action diagnosticity. In contrast, purported moderators and process measures of the dissonance account received no support when tested. Additionally, we systematically address potential confounds associated with varying levels of "noise" in preference expression through decisions, an issue that had encumbered previous paradigms on preference modulation. Our findings suggest that changes in preference induced by ordinary day-to-day decisions primarily stem from an ongoing process of information inference and updating, rather than dissonance reduction. This research also provides insights into the previously unforeseen consequences of framing interventions in policy and business.

Keywords: decision-induced preference modulation, choose/reject framing, self-perception theory, dissonance theory, self-preference inferences, action diagnosticity

Public Significance (260)

Decisions framed as opportunities to choose or reject are increasingly used by choice architects to influence decision outcomes in business and policy. However, since decisions can alter subsequent preferences (reviewed in Enisman et al., 2021), it is important to understand whether such framing interventions may inadvertently influence post-decision preferences and attitudes beyond their intended immediate impact.

This research examines *whether*, *how*, and *why* framing impacts post-decision consequences. We repeatedly show (in 13 total pre-registered experiments, $N = 9,927$ participants in North America and Asia) that framing a decision between attractive options as a reject action intensifies preferences more than framing the same decision as a choose action. This framing effect is observed across diverse stimuli (e.g., products, art prints, food, photos, and words), is mitigated when the options are similar in attributes, and is reversed in direction when the decision is between unattractive options. Our results corroborate an inference explanation for decision-induced preference modulation, based on key insights from self-perception theory (Bem, 1967) and the theory of representational exchange (Cushman 2020): individuals extract new information about their underlying preferences from observing their own decisions, and reject framing tends to render decisions more self-diagnostic than choose framing. Meanwhile, our results challenge the fit of a previously more popular dissonance explanation (Festinger, 1957) for preference modulation in ordinary day-to-day decisions. Therefore, our findings also suggest that preference modulation may be more commonplace than previously implied. Beyond theoretical contributions, this research also provide novel insights into how framing interventions can be leveraged toward positive outcomes in practical domains including marketing, management, and policymaking.

Framing Affects Post-Decision Preferences Through Self-Preference Inferences (and Probably not Dissonance)

The idea that decisions enhance subsequent preferences was initially proposed by some of the most influential theories in psychology (e.g., Festinger, 1957, 1962; Bem, 1967, 1972). Consider John, who is faced with a choice between two preferred art prints, a Monet and a Picasso. These theories predict that if John selects the Monet, this decision will intensify his preference for the Monet over the Picasso. This phenomenon, known as post-decision preference modulation, has captivated psychologists for generations and has now been empirically substantiated after decades of testing and corrections (see Enisman et al., 2021 for a review). The accumulating evidence supporting this phenomenon challenges a foundational assumption about the relationship between preferences and decisions—conventionally, preferences are assumed to influence decisions, but not the reverse.

While daily decisions are typically presented as opportunities to choose, they can also be presented as opportunities to reject. From customizing products to selecting medical treatments and planning career paths, most decisions can be framed either as choices of preferred options or rejections of non-preferred ones. The power of choose/reject framing in altering decision outcomes has been well documented (e.g., Dhar & Wertenbroch, 2000; Shafir, 1993). However, limited research has examined the potential impact of framing on post-decision consequences beyond the decision outcomes per se, whereas no research has delved into how framing might influence decision-induced preference modulation (or preference modulation, for short).

This gap in the literature presents a unique opportunity to advance both theory and practice. Theoretically, distinct explanations for the phenomenon have been offered by influential psychological theories including cognitive dissonance theory (Festinger, 1957)

and self-perception theory (Bem, 1967). Yet, research has not resolved the theoretical ambiguity regarding the phenomenon's primary psychological mechanism(s), despite significant recent progress in documenting the phenomenon and addressing previous methodological issues (see review in Enisman et al., 2021). To this end, examining how and when framing affects preference modulation may give rise to new opportunities to disambiguate mechanisms. Practically, the widespread adoption of framing techniques in business and policy (Thaler & Sunstein, 2008) also calls for research into their downstream consequences. A notable example is the increasing popularity of negative advertising strategies in political campaigns, which encourage voters to reject an alternative candidate as opposed to traditional positive advertising that promotes support for a focal candidate. It would be concerning if such a shift toward reject-framed strategies inadvertently exacerbates political polarization in society beyond merely impacting voting outcomes.

Motivated to address these questions, we examine whether, how, and why decision framing impacts preference modulation. The primary goal of this research is to test the effect of choose/reject framing on post-decision preference modulation. A secondary goal is to shed new light on its mechanisms toward addressing the tension between alternative psychological explanations.

We first draw on self-perception theory (Bem, 1967, 1972) and a more recent framework of representational exchange (Cushman, 2020) to introduce a general information-inference account. This inference account suggests that observing one's own decision prompts people to infer new information about their "true" preferences and update their preference expressions accordingly. Based on this account, we propose a series of hypotheses regarding how choose/reject framing will affect preference modulation. The main hypothesis is that reject framing, through heightening the perceived diagnosticity of a decision, will

induce more preference modulation than choose framing, resulting in a larger preference gap between the preferred option and the non-preferred option.

We then distinguish this inference account from the dissonance account offered by cognitive dissonance theory (Festinger, 1957, 1962). Dissonance theory posits that post-decision preference modulation is caused by an aversive state of dissonance and an accompanied motivation to reduce dissonance. Unlike the inference account, the dissonance account does not provide a clear *ex ante* prediction on the framing effect per se. Nevertheless, it furnishes flexible *ex post* alternative explanations for our hypothesized framing effect as well as for its opposite effect. Additionally, dissonance theory has articulated necessary antecedents for observing preference modulation that differ from those necessitated by the inference account. We will empirically test between these mechanisms in the framing context.

We test our hypotheses in seven pre-registered experiments ($N = 6,248$ participants from North America and Asia) and six pre-registered supplemental experiments ($N = 3,679$ participants from North America and Asia). We repeatedly find the proposed framing effect, the diagnosticity-based process, and the moderating role of attribute similarity and choice set valence on the framing effect, all consistent with predictions of the inference account. Meanwhile, our results do not support the process measures and moderators of the dissonance account.

Notably, in testing the framing effect, our experiments do not involve comparisons of pre- versus post- decision preferences (known as the “spreading of preferences” or “preference spread”), a procedure that had been shown to be susceptible to methodological confounds (see critique and proof in Chen and Risen 2010; see a review of rectified studies in Enisman et al., 2021). Instead, we solely measure and compare the relative strengths of post-

decision preferences between choose and reject frames. The original methodological confounds of “preference spread” are hence not relevant for testing this framing effect. Nonetheless, within the framing paradigm, we systematically assess whether observed differences in post-decision preferences were attributable to varying levels of “noise” in preference expression through (choose and reject) decisions, addressing potential confounds that are conceptually related to prior critiques.

Decision-Induced Preference Modulation

Why do decisions change subsequent preferences? Two of the most influential theories in social psychology have offered distinct answers to this question. One centers on an unpleasant state of dissonance induced by the decision (e.g., Festinger, 1957, 1962), whereas the other emphasizes a spontaneous information inference from the decision (e.g., Bem, 1967, 1972; Cushman, 2020). These theories broadly extend to the phenomenon of rationalization, the observation that after a person performs an action, their subsequent beliefs and attitudes tend to change in a manner that helps make the initial action appear more rational (see Cushman, 2020 for a review).

There is apparent tension between the dissonance account and the inference account. However, because they involve separate sets of psychological mechanisms, the two accounts typically do not generate “unequivocally conflicting predictions” (Greenwald, 1975). As such, a growing consensus over the past few decades suggests that the two accounts may offer complementary rather than competing explanations (Cushman, 2020; Fazio et al., 1977; Greenwald, 1975), and that they may apply primarily in different domains (as reviewed by Cushman, 2020 and Harmon-Jones & Mills, 1999). Indeed, whereas self-perception theory has enjoyed extensive support in contexts involving attitude-congruent behaviors, such as self-monitoring and self-signaling behaviors (e.g., Bodner & Prelec, 2002; Dhar &

Wertebroch, 2012; Rosenberg, 1979), evidence for dissonance has been documented primarily in contexts involving attitude-discrepant behaviors, such as attitude changes following obviously counter-attitudinal actions or unexpectedly painstaking efforts (Elliot & Devine, 1994; Snyder & Ebbesen, 1972).

Then, which account better explains the observation of preference modulation in ordinary day-to-day decisions, such as in the example of John deciding between art prints? A brief glance through the literature will reveal numerous casual references to such examples as a “dissonance phenomenon,” implying dissonance reduction as its primary explanation. However, this assumption has never been empirically assessed. In fact, leading scholars in this area have explicitly noted that existing research has never tested between the two explanations in the context of decision-induced preference modulation (e.g., Footnote 1 in Chen & Risen, 2010, p. 573; Footnote 4 in Risen and Chen, 2010, p. 1161; see also Enisman et al., 2021, p. 24). As of today, the answer to this question remains unclear.

This unresolved question regarding the *psychological mechanism(s)* underlying preference modulation should not be conflated with previous contentions regarding the *validity* of evidence for preference modulation. In 2010, Chen and Risen (2010) identified critical confounds in the free-choice paradigm (FCP; Brehm, 1956) that had been the dominant paradigm used to study preference modulation for more than half a century. Since then, researchers used modified paradigms to correct for the confound and have re-established robust evidence for preference modulation (reviewed in Enisman et al., 2021).

Therefore, the key question in this research pertains to *why* decisions alter subsequent preferences, not *whether* they do.¹

Next, we consider the potential impact of (choose vs. reject) framing on preference modulation based on each theoretical perspective. We begin by introducing the inference account, which generates a series of straightforward hypotheses on a framing effect and its primary characteristics. Subsequently, we assess these hypotheses from the dissonance perspective and discuss their differences. As will be seen, these two accounts do not make identical predictions on the presence, direction, and boundaries of the framing effect. These differences will enable us to disentangle the two accounts both conceptually and empirically.

The Inference Account

The gist of the information-inference account is that people glean insights into an actor's underlying preferences through their actions. Extensive research in social psychology has established a ubiquitous process of attributional inference—people learn about others' preferences by observing others' actions and attributing them to internal or external factors (Jones & Davis, 1965; Kelley, 1967). Self-perception theory (Bem, 1967, 1972) builds on this premise and takes one step further to contend that people, lacking perfect insight into their own preferences, may also infer their own preferences through the observation of their actions (Bem, 1972, p.2).

Extending this key insight, a recently proposed and well-received framework of representational exchange (Cushman, 2020) posits that inferring preferences from one's own actions is not only plausible, but also broadly adaptive. Synthesizing over a century of

¹ In fact, if choose/reject framing induces different degrees of preference modulation, then preference modulation must be true. Hence, our empirical tests for the framing effect may serve as an independent validation for the existence of preference modulation without having to examine the “preference spread” per se.

behavioral, developmental, and neurocognitive evidence on the multiple processes underlying actions and decisions (e.g., Dolan & Dayan, 2013; Kahneman, 2011; Thorndike, 1898), Cushman (2020) suggests that one's actions may reveal useful information about their underlying preferences that are tied to their instincts, reflexes, and habits, which would be largely inaccessible via introspection and deliberation (see also Nisbett & Wilson, 1977). Hence, preference modulation can be viewed as the result of new preference information—which are revealed by the observable decision—being integrated into the mental representations of one's own preferences. More concretely, in the earlier example with John, this inference account suggests that John's decision of acquiring the Monet will prompt him to infer that he probably liked the Monet more than the Picasso, which then influences how he rates the two paintings subsequently.

Critically, not all actions are equally diagnostic of the actor's underlying preferences. In social attribution, negative actions (e.g., violating a social norm or writing negative reviews) have a larger impact than positive actions (e.g., conforming to a social norm or writing positive reviews) on observers' inferences about the actor's dispositional preference (Skowronski & Carlston, 1989; Mizerski, 1982). These findings have been attributed to multiple psychological mechanisms, one of which is the greater attentional resources evoked by the negative implications of actions than positive ones (see Baumeister et al., 2001 for a review). Importantly, the reject framing emphasizes the negative aspects of a decision (e.g., potential losses) whereas the choose framing emphasizes its positive aspects (e.g., potential gains; Levin et al., 1998; Houston et al., 1991; Shafir, 1993; Tversky & Kahneman, 1981). Therefore, if observers tend to extract more information about another actor's dispositional preference from negative actions than positive actions, then they may likewise deem their

own reject actions as more diagnostic about their own underlying preferences than choose actions, all else equal.

Indeed, lending support to the different impact of choose/reject framing on the process of decision-making, converging evidence shows that a reject decision tends to evoke a more deliberative judgment process than an otherwise equivalent choose decision (Sokolova & Krishna, 2016; Krishnamurthy & Nagpal 2008; Laran & Wilcox, 2011). Therefore, based on existing associations between the depth of information processing and decision frames, we propose that people may treat their own reject actions as more informative about their underlying preference than their choose actions—as if reject actions tap “deeper” into their true preferences than choose actions. Subsequently, when extracted preference information is integrated into one’s existing knowledge and belief about own preferences², greater preference modulation should follow a reject-framed decision than a choose-framed decision.

In summary, based on the inference account, we hypothesize that, because reject actions are deemed more diagnostic than choose actions about one’s underlying preferences, reject framing may intensify post-decision preferences more than choose framing. Hence we have H1 and H2.

Hypothesis 1 (a framing effect): Reject framing will lead to greater post-decision preference modulation (i.e., a larger post-decision preference gap between the preferred and non-preferred alternatives) than choose framing.

² Bem (1972) originally proposed the preference updating is an updating of one’s belief about own preferences, whereas Cushman (2020)’s theorizing is decidedly less skeptical, arguing for an updating of actual preferences. This theoretical distinction between them does not affect our hypotheses in the current research.

Hypothesis 2 (on action diagnosticity): The reject-framed decision will be perceived by decision-makers as more diagnostic about their own underlying preferences than the choose-framed decision, all else equal.

The Dissonance Account

Cognitive dissonance theory (Festinger, 1957, 1962) explains preference modulation as the result of an attempt to reduce decision-induced dissonance. Dissonance, an aversive emotional-motivational state, is theorized to occur when there is inconsistency between cognitions, such as a discrepancy between preferences and decisions. People are purportedly motivated to reduce dissonance, such as by revising their post-action preferences to be more closely aligned with the preceding decision. Classic demonstrations of dissonance include individuals' evaluation of a boring task being elevated after having complied to tell others that the task was enjoyable (Festinger & Carlsmith, 1959). In such contexts, evidence for dissonance has been obtained through measures of self-reported psychological discomfort (Elliot & Devine, 1994) as well as physiological arousal (Croyle & Cooper 1983).

More concretely, for John in our example, dissonance theory suggests that if John finds the two art prints equally attractive yet decided to acquire the Monet only, the partial conflict between his preference and his decision would produce dissonance. Consequently, John would increase his liking of the Monet over the Picasso to alleviate dissonance. This explanation is also plausible, while certainly distinct from the inference explanation.

Now, in the framing context, we have proposed a larger post-decision preference gap in the reject frame than in the choose frame, based on the inference account. To the best of our knowledge, the dissonance account does not yield a clear *ex ante* prediction on the

presence or the direction of such a framing effect.³ Nonetheless, it can flexibly furnish *ex post* explanations for a framing effect to be observed in either direction. That is, dissonance theory remains a plausible explanation as long as *greater* dissonance is observed in the frame where a larger preference gap is found, *and if* this difference in dissonance accounts for at least some of the differences in preference gaps between frames.

These are empirically testable. Therefore, we will collect measures of both the perceived diagnosticity of the decision (per the inference account) and the experience of dissonance induced by the decision (per the dissonance account). We will then examine if any observed differences in preference modulation are attributable to these process measures, respectively. These tests will provide the first set of evidence to assess the relative fit of theories.

Next, we again use the inference account to derive key theoretical moderators of the framing effect: option similarity and choice-set valence. As will be seen, the two accounts' ranges of applicability will be further differentiated by the potential boundary conditions.

Option Similarity

From the perspective of information inferences, the degree to which a decision induces subsequent preference modulation is inherently constrained by the amount of

³ The total magnitude of dissonance is believed to be determined by the relative strengths of dissonant versus consonant cognitions (Brehm, 1956; Festinger & Carlsmith, 1959). Framing can impact both types of cognitions. For example, the reject frame often draws more attentional resources to information processing than the choose frame (Sokolova & Krishna, 2016), potentially heightening the conflict between action and attitude (e.g., between “rejecting the Picasso” and “this Picasso painting is cool”). Meanwhile, the reject frame may focus more attentional resources on negative attributes than the choose frame (Houston et al., 1991; Shafir, 1993), potentially strengthening the consonance between the reject action and the negative attributes of the rejected option (e.g., “this Picasso painting uses dull colors”). Without being able to quantify the relative magnitude of these various factors, it is infeasible to predict how overall dissonance may differ between frames.

information extractable from the decision. When options in the choice set share more identical attributes, less new information can be inferred from the decision. In essence, as the similarity of attributes between options increases, the total preference information available decreases. Consistent with this logic, studies have shown that choices made from options with similar self-control connotations (e.g., “vice” versus “vice”) are perceived as less diagnostic of the decision-maker’s dispositional self-control than choices made from distinct options (e.g., “vice” versus “virtue”; Dhar & Wertenbroch, 2012). When little new information about the decision-maker’s preference can be gleaned from a decision, preferences tend to remain relatively unchanged before and after the decision. This limits the impact of decision framing on preference modulation. Therefore, it follows that, a high similarity between options serves as a natural boundary condition for the framing effect.

Hypothesis 3 (option similarity as a moderator): The framing effect will attenuate when the options in the choice set are highly similar in attributes.

In agreement with the above reasoning, dissonance theory has also suggested that attribute dissimilarity between options is a necessary antecedent for decision-induced dissonance and hence preference modulation (Brehm & Cohen, 1959; Festinger, 1957). However, dissonance theory has specified *one more* necessary antecedent for dissonance—equal attractiveness of options. In fact, since the inception of dissonance theory, dissonance researchers have repeatedly stressed that dissonance emerges from decisions between options perceived as “*equally attractive*”; deciding between options that obviously differ in attractiveness is considered easy and hence unlikely to provoke internal conflict (e.g., Brehm 1956; Festinger, 1962; Shultz et al., 1999). This is why the free choice paradigm (FCP) always mandates equating choice alternatives in attractiveness in its first step (Brehm, 1956). Accordingly, dissonance researchers have reported in several FCP studies where preference

modulation no longer occurred after a decision between options of unequal attractiveness (e.g., Brehm & Cohen, 1959; Harmon-Jones & Harmon-Jones, 2002; Shultz et al., 1999).⁴

In summary, according to dissonance theory, two prerequisites must be met for dissonance-based preference modulation to occur: options must be *similar* in attractiveness but *dissimilar* in attributes. In contrast, the inference account only requires that options be *dissimilar* in attributes. Therefore, if we still observe the framing effect with options that are *dissimilar* in attractiveness, it would be challenging for the dissonance account to explain it.

Valence of Choice Set

Thus far, our theorizing has focused on decisions involving primarily attractive options. Notably, many framing effects initially established with attractive options are reversed when tested with unattractive options instead (e.g., Kessler et al., 1996; Meloy & Russo, 2004; Nagpal & Krishnamurthy, 2008; Perfecto et al., 2017; Wedell 1997). This reversal pattern has been attributed to the compatibility between the valence of the action and the valence of the choice set, which can affect the metacognitive fluency of information processing (Perfecto et al., 2017; Sokolova & Krishna, 2016; Wedell, 1997). Specifically, the act of choosing is deemed relatively more compatible with attractive options whereas the act

⁴ These reported findings should be interpreted with caution due to their faulty conceptual and empirical assumptions that have been discussed in an influential critique (Chen & Risen, 2010). First, Chen and Risen identified fatal methodological confounds arising from selection issues in the FCP protocol used in all these studies. Chen and Risen explicitly specified that these confounds undermine the validity of various claims in the quoted studies, including the role of “equal attractiveness” (see this critique on p. 587 and the mathematical proof on p.593, Chen and Risen 2010; see also Izuma and Murayama, 2013, p. 8). Second, none of the quoted studies actually measured dissonance. Instead, their conclusions about the characteristics of dissonance were based entirely on the implicit assumption that dissonance caused preference modulation; yet that assumption has never been empirically verified (see this critique in Footnote 1 in Chen & Risen, 2010, p. 573 and Footnote 4 in Risen and Chen, 2010, p. 1161), as mentioned earlier.

of rejecting is deemed relatively more compatible with unattractive options. In both cases, compatibility increases processing fluency, which can modulate subsequent outcomes, such as enhancing post-decision confidence (Perfecto et al., 2017; see also Meloy & Russo, 2004).

The critical role of metacognitive fluency in information processing suggests that the framing effect, as proposed based on the inference account, may also depend on the valence of the choice set. When the decision involves primarily attractive options, we have suggested that people read “deeper” into the reject action as if it reveals more underlying preferences than a comparable choose action. These associations between actions and preference strengths should be weakened or even reversed when decisions involve less attractive options, particularly between strongly unattractive options, where the reject action becomes more natural and prevalent (Buiten & Keren, 2009). Ultimately, it is conceivable that choosing between unattractive options may signal even stronger underlying preferences than rejecting between unattractive options.

Hypothesis 4 (choice-set valence as a moderator): The framing effect should be mitigated, and potentially reversed, when the choice set comprises less attractive alternatives.

Once more, the dissonance account may again provide a plausible *ex post* explanation of H4. However, that requires additional speculations about how compatibility and fluency affected different components of dissonance (see Footnote 2), and involves various empirical questions, some of which will be addressed as we test our hypotheses in order.

Experimental Paradigm

In each experiment, we randomly assigned participants to one of two framing (choose vs. reject) conditions. Participants were presented with the same binary choice-set and were asked to either choose a preferred option or reject a less preferred option, corresponding to

the assigned condition. After the decision, participants were asked to rate their preference for each option on a numerical scale. Our key dependent variable is the post-decision preference gap (Δ), computed by subtracting the rating of the unwanted (i.e., unchosen/rejected) option from the rating of the wanted (i.e., chosen/retained) option. All our studies used hypothetical scenarios as preference modulation was shown to occur similarly in hypothetical and real choices (e.g., Sharot et al., 2010).

As indicated in H1, we expect a larger post-decision preference gap in the reject condition than the choose condition. However, to interpret this expected difference as different degrees of preference modulation, we must rule out other factors that could speciously produce the same result, including random error and systematic error. Importantly, both decisions and preference ratings are “noisy” expressions of underlying preferences (Chen and Risen 2010). In fact, instances of self-contradictory decisions and preferences are inevitable in psychological experiments, stemming from various sources including vague initial preferences, changes of mind, miscomprehensions, inattention, and so on. These instances can lead to critical confounds if not carefully controlled.

Within our framing paradigm, two sources of such “noise” are worthy of attention. First, it is possible that the two groups of participants differ in their preference volatility or attentiveness at the outset of the experiment. Although such randomization failures can lead to spurious results (i.e., Type I error), their likelihood can be significantly reduced by established practices, such as true random assignment, large sample sizes, and repeated replications—all of which we closely adhere to.

The second source lies in how the imposed frames might differently affect the likelihood that decisions misrepresent existing preferences during the experiment. For example, if a participant chooses A but rates B higher, their calculated preference gap will be

negative in value (< 0), thereby shrinking the average gap in their condition. If one condition had significantly more of such participants than the other condition, we may observe the predicted difference in preference gap even if framing had no effect on preference modulation. To be clear, this possibility does not undermine the empirical validity of the proposed framing effect, nor its novelty. Rather, it suggests yet another psychologically meaningful alternative explanation: framing may affect the likelihood that a decision *misrepresents* existing preferences. This possibility is conceptually distinct from our proposal that choose and reject decisions *change* subsequent preferences to different degrees. Therefore, a clear theoretical interpretation of the observed effect necessitates that we address whether and to what extent the observed effect is attributable to such a possibility.⁵

To address these possibilities, we use two methods to systematically assess the impact of self-contradictory decisions and preferences on the framing effect. First, in each study, we compare the frequency of decision-inconsistent preferences between conditions, and we report the framing effect both with and without these instances included. (The rationale and procedure of this method are detailed in Study 1.) As can be seen in Table 1, few participants exhibited decision-inconsistent preferences (i.e., “DIP”), and the framing effect was highly similar with or without applying the additional “DIP” screener in all studies. These results demonstrate the minimal impact of decision-inconsistent preferences on the observed effect. Second, in Study 5, we adapted a previously tested method to fit the framing paradigm and confirm that potential confounds do not drive the framing effect.

⁵ While this possibility may be reminiscent of the famous confounds in the FCP, the original selection issues in the FCP are procedurally irrelevant for testing the framing effect in this research. Again, this is because we examine how framing impacts post-decision preference strengths alone without measuring any pre-decision preferences, or having to measure any.

Overview of Studies

We present seven experiments ($N = 6,248$ participants from North America and Asia) in the main text of this paper and six supplementary experiments ($N = 3,679$) in the SOM.

We first tested how decision framing affects preference modulation with a purchase scenario in Study 1 (H1). We then replicated this effect in Studies 2a and 2b while examining process variables to test between two theoretical accounts (H1 & H2). Next, we examined the role of option similarity (H3) in Study 3, and the role of choice-set valence (H4) in Studies 4a and 4b. Finally, we conducted Study 5 to hone in on the minimal impact of potential confounds on the framing effect.

Transparency and Openness

We pre-registered all experiments and report all manipulations, measures, and analyses. We share all data, unabridged survey materials, analysis code, and the SOM on OSF (<https://osf.io/qhcdy/>). We report all primary results in the main text. We summarize secondary analyses and results in the main text and report their full details in the SOM.

Sampling

Our target sample sizes followed the heuristic of at least 125 participants per between-subjects condition. Sample sizes were doubled in replication studies when resources allowed (Studies 1 & 2b), and at least tripled in moderation studies (Studies 4a, 4b & 5) except for Study 3, in which the sample size was capped the size of a subject pool.

Screening

We used the same standard screening procedure in all studies: excluding all incomplete responses and duplicate IP addresses in online samples (and excluding all incomplete responses for the subject-pool sample). We also excluded responses that failed an instructional manipulation check (IMC; Oppenheimer et al., 2009) during periods when

online platforms reported survey bots. All analyses were conducted upon completion of standard screening. These standard screening criteria were pre-registered, and they excluded similar numbers of participants from the choose and reject conditions in all studies (Table S1.1, SOM). Table 1 summarizes the main results on preference modulation. The “Before DIP Screening” columns directly correspond to the pre-registered analyses. The DIP screening method, used for post-hoc robustness tests, was not pre-registered nor intended as part of standard screening. The DIP-screened results (in the three rightmost columns) remain similar.

Table 1. The framing effect was observed in all the test (non-shaded) conditions and attenuated or reversed in the proposed boundary (shaded) conditions. The results are similar before and after DIP screening.

Study	Post-decision preference gap (Δ) and cell size (n)	Before DIP screening			After DIP screening		
		Choose Frame	Reject Frame	Framing Effect	Choose Frame	Reject Frame	Framing Effect
1	Δ	2.05 [1.80, 2.31]	2.56 [2.30, 2.81]	$t(526) = 2.71$, $p = .007$, $d = 0.24$, [0.07, 0.41]	2.12 [1.87, 2.38]	2.64 [2.39, 2.89]	$t(516) = 2.85$, $p = .005$, $d = 0.25$, [0.08, 0.42]
	n	262	266		257	261	
2a	Δ	2.08 [1.67, 2.48]	3.02 [2.60, 3.43]	$t(254) = 3.18$, $p = .002$, $d = 0.40$, [0.15, 0.65]	2.18 [1.80, 2.56]	3.26 [2.87, 3.64]	$t(246) = 3.91$, $p < .001$, $d = 0.50$, [0.25, 0.75]
	n	130	126		127	121	
2b	Δ	2.90 [2.66, 3.14]	3.36 [3.12, 3.59]	$t(629) = 2.67$, $p = .008$, $d = 0.21$, [0.06, 0.37]	3.06 [2.84, 3.28]	3.45 [3.23, 3.67]	$t(615) = 2.47$, $p = .014$, $d = 0.21$, [0.04, 0.36]
	n	317	314		308	309	
3 (Moderately Similar Options)	Δ	1.74 [1.45, 2.04]	2.34 [2.03, 2.64]	$t(231) = 2.75$, $p = .006$, $d = 0.36$, [0.10, 0.62]	1.85 [1.57, 2.14]	2.40 [2.10, 2.69]	$t(225) = 2.59$, $p = .010$, $d = 0.35$, [0.09, 0.61]
	n	120	113		116	111	
3 (Highly Similar Options)	Δ	1.49 [1.25, 1.72]	1.48 [1.25, 1.70]	$t(235) = -0.06$, $p = .95$, $d = -0.01$, [-0.26, 0.25]	1.49 [1.25, 1.72]	1.50 [1.27, 1.72]	$t(234) = 0.06$, $p = .96$, $d = 0.01$, [-0.25, 0.26]
	n	113	124		113	123	
4a (Positive Choice Set)	Δ	4.09 [3.64, 4.54]	5.32 [4.86, 5.78]	$t(675) = 4.20$, $p < .001$, $d = 0.32$, [0.17, 0.48]	4.25 [3.83, 4.67]	5.52 [5.09, 5.94]	$t(660) = 4.50$, $p < .001$, $d = 0.35$, [0.20, 0.50]
	n	344	333		336	326	
4a (Negative Choice Set)	Δ	5.53 [5.08, 5.98]	5.44 [4.98, 5.90]	$t(681) = -0.25$, $p = .80$, $d = -0.02$, [-0.17, 0.13]	5.76 [5.34, 6.18]	5.96 [5.53, 6.39]	$t(660) = 0.62$, $p = .54$, $d = 0.04$, [-0.11, 0.20]
	n	346	337		338	324	

4b (Positive Choice Set)	d	3.04 [2.77, 3.30]	4.46 [4.19, 4.73]	$t(784) = 7.39,$ $p < .001,$ $d = 0.52, [0.39, 0.67]$	3.19 [2.92, 3.47]	4.62 [4.34, 4.89]	$t(724) = 7.20,$ $p < .001,$ $d = 0.54, [0.39, 0.69]$
	n	398	388		369	357	
4b (Negative Choice Set)	d	3.67 [3.44, 3.89]	3.39 [3.16, 3.62]	$t(783) = -1.66,$ $p = .097,$ $d = -0.12, [-0.26, 0.02]$	4.09 [3.86, 4.32]	3.58 [3.36, 3.81]	$t(685) = -3.11,$ $p = .002,$ $d = -0.24, [-0.39, -0.09]$
	n	397	388		337	350	
5 (Decide- Rate)	d	1.62 [1.44, 1.80]	2.49 [2.31, 2.67]	$t(714) = 6.68,$ $p < .001,$ $d = 0.50, [0.35, 0.65]$	1.69 [1.52, 1.87]	2.52 [2.35, 2.70]	$t(703) = 6.52,$ $p < .001,$ $d = 0.49, [0.34, 0.64]$
	n	356	360		348	357	
5 (Rate- Decide)	d	1.60 [1.40, 1.81]	1.39 [1.18, 1.59]	$t(714) = -1.46,$ $p = .15,$ $d = -0.11, [-0.25, 0.04]$	1.73 [1.52, 1.93]	1.52 [1.31, 1.72]	$t(684) = -1.44,$ $p = .15,$ $d = -0.12, [-0.26, 0.03]$
	n	356	360		342	344	

Study 1—A Purchase Decision

Study 1 tested the framing effect in a purchase scenario, pre-registered at

<https://aspredicted.org/blind.php?x=gp5a42> .

Method

Participants. We recruited 550 Prolific participants and received 568 data entries.

After standard screening (see full screening details in Table S1.1, SOM), we obtained 528 valid responses from 230 male and 298 female participants with an average age of 34.

Procedure. Participants were asked to imagine making a purchase decision between two preferred smartphone cases, both compatible with their current phone. The two options were both in black and white, from the same brand and priced the same, presented side by side (Figure 1). The only difference was in their patterns: checkered versus striped.

Figure 1. The two smartphone cases used in Study 1 and Study 2a.



We randomly assigned participants to the two framing (choose vs. reject) conditions. In the choose condition, participants were instructed to drag their preferred option into a box labeled “I choose...” in response to the question “Which phone case would you choose?”. In the reject condition, participants were instructed to drag their non-preferred option into a box labeled “I reject...” in response to the question “Which phone case would you reject?”. Then participants were asked to indicate how much they valued each option on a 10-point scale from 1 (*not at all*) to 10 (*very much*). Participants first rated the wanted (i.e., chosen/retained) option, and then rated the unwanted (i.e., non-chosen/rejected) option on a separate page. (Further procedural details for the framing manipulations and preference measures in all studies are summarized in Table S1.2 in the SOM.)

Last, participants indicated their gender and age. Participants were asked to select between “male” and “female” options in Studies 1 and 2a. A third “non-binary/others” option was added in all subsequent studies.

Results

The post-decision preference gap in the reject condition ($\Delta_{\text{reject}} = 2.56$, 95% CI = [2.30, 2.81]) was larger than that in the choose condition ($\Delta_{\text{choose}} = 2.05$, 95% CI = [1.80, 2.31]), $t(526) = 2.71$, $p = .007$, $d = 0.24$, 95% CI = [0.07, 0.41]; see also Table 1). Both options received moderately positive ratings (see Table 2).

Table 2. Preference ratings by option and by frame in Study 1.

	<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
	Choose Frame	Reject Frame	Choose Frame	Reject Frame
Wanted Option	5.02 [4.71, 5.32]	5.42 [5.12, 5.72]	5.00 [4.70, 5.31]	5.42 [5.12, 5.73]
Unwanted Option	2.96 [2.71, 3.21]	2.86 [2.61, 3.11]	2.88 [2.64, 3.12]	2.78 [2.54, 3.02]

DIP Screening. We conducted an additional step of analysis to assess whether the larger gap in the reject condition incidentally resulted from unequal distributions of decision-inconsistent preferences (DIP) between conditions. To illustrate the rationale of this method, consider two participants: Ann, who rated her wanted option at least as positively as her unwanted option, and Ed, who rated his wanted option less positively than his unwanted option. Most participants are Anns, and few are Eds. Ann's decision and preference ratings were internally consistent, whereas Ed's were internally inconsistent—his decision and preference ratings formed a preference reversal.

As aforementioned, Ed's preference reversal could result from various sources of noise. Critically, if there were more Eds in one condition than the other, the calculated preference gaps could be skewed, inflating their difference between conditions. More specifically, this is because, in calculating the post-decision preference gap for any participant, the minuend is always the “wanted” option and the subtrahend is always the “unwanted” option, which are determined by the decision (and not by the preference ratings). Therefore, Ann's post-decision preference gap will be ≥ 0 , whereas Ed's will be < 0 . This means each Ed will shrink the average gap in his condition, and how much he shrinks the gap depends on the extremeness of his ratings. For simplicity, if we assume that each Ed's ratings were identically extreme. Then, if more Eds emerged in the choose condition, then including all Eds in the data would inflate the average gap in the reject condition relative to the choose condition. Alternatively, if more Eds emerged in the reject condition, then excluding all Eds from the data would instead inflate the average gap in the reject condition relative to the choose condition.

Therefore, to assess whether the larger preference gap observed in the reject condition was inadvertently inflated by unequal distributions of Eds' between conditions, we compare

the framing effect before and after excluding Eds in the data, namely, DIP screening. If we find similar effects before and after DIP screening, then the framing effect is unlikely a result of inflated gap calculations. However, if we find a significant framing effect only before or only after DIP screening, then further tests should be conducted to evaluate the extent to which decision-caused preference modulation accounted for the effect, as opposed to Eds' differential presence between frames. As such, the DIP method errs on the side of false negatives (i.e., Type II error), which is appropriate for its primary purpose of monitoring potential confounds for further testing. (We will discuss the uses and limitations of the DIP method more thoroughly in the General Discussion.)

In Study 1, DIP screening excluded similar ratios of participants in both conditions (2% vs. 2%), $\chi^2(1) = .001, p = .98$. The framing effect persisted with a similar magnitude after DIP screening, $t(516) = 2.85, p = .005, d = 0.25, 95\% \text{ CI} = [0.08, 0.42]$ (see also Tables 1 & 2). Therefore, the effect is not attributable to potential calculation confounds.

Discussion

Study 1 documents initial evidence for the hypothesized framing effect: reject framing intensified subsequent preferences more than choose framing. Since very few participants had preference reversals in either condition, their presence or absence barely affected the effect. Thus, the framing effect is unlikely to be explained by unequal distributions of decision-preference inconsistencies and associated alternative explanations (e.g., the different likelihood that choose and reject decisions misrepresent existing preferences). Instead, the framing effect in the current study primarily reflects greater preference modulation caused by the reject decision than the choose decision.

Some might wonder, did the preference gap differences between frames emerge primarily around the wanted option or around the unwanted option? Although our theorizing

does not yield a prediction on this second-order question, when we compared their differences, we found that both contributed to the gap across all studies. These additional results are summarized in the General Discussion with full details reported in Table S4.1.

Currently, the results in Study 1 are compatible with both inference and dissonance accounts. We start to disentangle these two accounts in subsequent studies, all of which include the DIP screening method to monitor potential confounds.

Studies 2a and 2b—Diagnosticity or Dissonance?

We tested both H1 and H2 in these two studies. To reiterate, the inference account suggests that decision-makers perceive their own reject action as more diagnostic of their underlying preferences than choose action, which results in more preference updating. The dissonance account suggests instead that if more preference updating is found in the reject condition, then it should be attributed to a stronger dissonance experienced by the decision-makers that is induced by the reject action than the choose action. Correspondingly, we measured both perceived action diagnosticity and experienced dissonance, as exploratory tests in Study 2a, and then confirmed in Study 2b.

Study 2a

We pre-registered this study at <https://aspredicted.org/blind.php?x=mk9ag5> .

Method

Participants. We recruited 280 MTurk participants and received 307 data entries. After standard screening, we obtained 256 valid responses from 163 male and 93 female participants with an average age of 36.

Procedure. Study 2a was based on the same stimuli and procedure in Study 1. One minor difference from Study 1 was that participants were first asked to rate the

chosen/rejected option, followed by the unchosen/retained option. All subsequent studies measured preference ratings in this order.

After those, we included three sets of pre-registered exploratory variables. First and most pertinent to our theorization was a 3-item measure of perceived action diagnosticity, adapted from Andersen (1984) and Touré-Tillery & Light (2018). Participants were asked to indicate the degree to which they agreed that their action of [choosing/not choosing/retaining/rejecting] each option “reflects my real preference,” “... is an opportunity for me to express myself,” and “... is a reflection of my identity,” from 1 (*definitely disagree*) to 10 (*definitely agree*).

Second, we included a 3-item measure of experienced dissonance, adapted from prior research (Elliot & Devine, 1994; Gerard, 1967; Izuma et al., 2010). Participants were asked to indicate how “difficult,” “conflicted,” and “uncomfortable” it was to [choose/not choose/retain/reject] each option from 1 (*not at all*) to 10 (*very*).

Third, we asked participants how definite their actions of [choosing/not choosing/retaining/rejecting] the options felt on a 10-point scale from 1 (*not at all*) to 10 (*very*). Based on prior findings that post-decision satisfaction depends on perceived decision closure (Gu et al., 2013), we included this measure to explore the possibility that preference updating, if attributable to preference inferences from own decisions, may also be facilitated when a stronger sense of closure is perceived from the decision.

Last, we included two sets of individual difference measures to explore if the framing effect was contingent on any personal traits, such as decisiveness (measured with the “trait decisiveness” subscale from the Need for Cognitive Closure Inventory in Webster & Kruglanski, 1994) and self-esteem (Rosenberg, 1979).

Results

The post-decision preference gap in the reject condition ($\Delta_{\text{reject}} = 3.02$, 95% CI = [2.60, 3.43]) was larger than that in the choose condition ($\Delta_{\text{choose}} = 2.08$, 95% CI = [1.67, 2.48]), $t(254) = 3.18$, $p = .002$, $d = 0.40$, 95% CI = [0.15, 0.65] (Tables 1 & 3), replicating the framing effect.

Table 3. Preference ratings by option and by frame in Study 2a.

	<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
	Choose Frame	Reject Frame	Choose Frame	Reject Frame
Wanted Option	5.54 [5.09, 5.99]	5.83 [5.37, 6.28]	5.58 [5.13, 6.04]	5.91 [5.45, 6.37]
Unwanted Option	3.46 [3.09, 3.84]	2.81 [2.43, 3.19]	3.40 [3.03, 3.77]	2.65 [2.27, 3.03]

DIP screening resulted in similar additional attritions in the two conditions (2% vs. 4%), $\chi^2(1, N = 256) = 0.58$, $p = .45$, which did not significantly influence the results (Tables 1 & 3). Next, we report the process findings based on data before DIP screening. Results after DIP screening remain highly similar and are reported in the SOM (section 5).

Action Diagnosticity (Cronbach's $\alpha = .85$). Consistent with H2, participants in the reject condition perceived their decision to be more diagnostic of own preferences than participants in the choose condition ($M_{\text{reject}} = 6.12$, 95% CI = [5.73, 6.51] vs. $M_{\text{choose}} = 5.56$, 95% CI = [5.18, 5.94], $t(254) = 2.05$, $p = .042$, $d = 0.25$, 95% CI = [0.01, 0.50]). Action diagnosticity positively predicted the size of post-decision preference gap ($b_{\text{diagnosticity}} = 0.50$, $SE = 0.06$, $p < .001$) and reduced the coefficient of framing on the preference gap when included in the regression model (from $b = 0.94$, $p = .002$ to $b = 0.66$, $p = .013$). Action diagnosticity mediated the framing effect (indirect effect = -0.28, 95% CI = [-0.57, -0.02]).

Experienced dissonance ($\alpha = .91$). Participants in the reject condition reported less dissonance than participants in the choose condition ($M_{\text{reject}} = 2.14$, 95% CI = [1.82, 2.46] vs.

$M_{\text{choose}} = 2.62$, 95% CI = [2.31, 2.94], $t(254) = 2.11$, $p = .036$, $d = 0.26$, 95% CI = [0.02, 0.51]). This is the opposite of what would be expected if dissonance explained the observed framing effect.

Decision closure. Perceived decision closure moderated the framing effect (interaction $F(1, 252) = 22.18$, $p < .001$, $\eta_p^2 = .08$; $M_{\text{closure}} = 7.79$, J-N value = 7.57), with a relatively stronger effect among participants who perceived greater decision closure (conditional effect at $M_{\text{closure}} - 1\text{SD} = 0.60$, 95% CI = [-0.14, 1.34]; conditional effect at $M_{\text{closure}} + 1\text{SD} = -1.91$, 95% CI = [-2.64, -1.18]). These results suggest that, in this study, perceived decision closure amplified post-decision preference updating—preference updating was more likely to follow a decision that felt completed.

PID measures. Neither trait decisiveness nor self-esteem moderated the framing effect ($F_s < 1.57$, $p_s > .21$; see details in section 6 in the SOM). In other words, the framing effect applied to participants irrespective of individual differences in these traits.

Study 2b

We conducted Study 2b to further replicate the framing effect with different stimuli, and to provide a confirmatory test for the process findings in Study 2a. The study was pre-registered at https://aspredicted.org/9L9_V56.

Method

Participants. We pre-registered the Pocock boundary method (Pocock, 1977) for the framing effect on action diagnosticity, the effect size of which was smaller than that of the preference gaps in Study 2a. This ensured sufficient power to detect both effects. More specifically, we planned to recruit 626 participants on Prolific for at least 80% power to detect a between-participants difference in action diagnosticity with a significance level of $p < .05$ (based on the effect size from previously conducted studies using similar stimuli; see

Studies S5 & S6). We specified that if the above analysis reached the Pocock boundary of $p < .0294$ (Pocock, 1977), then we would terminate data collection; if not, then we would continue data collection until reaching a total sample size of 1043 participants for 95% power on the same analysis.

We received 667 data entries and obtained 631 valid responses ($M_{\text{age}} = 39$, 327 male, 292 female, 12 non-binary/others) after standard screening. The effect of framing on action diagnosticity reached $p = .022$. Hence, we terminated data collection.

Procedure. We kept all procedures in Study 2b identical to those in Study 2a except the stimuli. Participants were presented with two art prints: “The Dessert” (1901) by Pablo Picasso, and “Morning on the Seine Near Giverny” (1897) by Claude Monet. The paintings were presented without the titles or names of the artists (Figure 2).

Figure 2. The two paintings in Study 2b.



Results

The post-decision preference gap in the reject condition ($\Delta_{\text{reject}} = 3.36$, 95% CI = [3.12, 3.59]) was larger than that in the choose condition ($\Delta_{\text{choose}} = 2.90$, 95% CI = [2.66, 3.14]), $t(629) = 2.67$, $p = .008$, $d = 0.21$, 95% CI = [0.06, 0.37]). Again, this framing effect was observed both with and without DIP screening (Tables 1 & 4). DIP screening resulted in similar attritions in the two conditions (3% vs. 2%), $\chi^2(1, N = 631) = 1.13$, $p = .29$, which did not affect the results.

Table 4. Preference ratings by option and by frame in Study 2b.

	<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
	Choose Frame	Reject Frame	Choose Frame	Reject Frame
Wanted Option	7.39 [7.21, 7.58]	7.83 [7.64, 8.02]	7.46 [7.28, 7.64]	7.89 [7.71, 8.07]
Unwanted Option	4.49 [4.26, 4.72]	4.47 [4.24, 4.70]	4.40 [4.17, 4.63]	4.44 [4.21, 4.67]

Action Diagnosticity ($\alpha = .81$). Again, perceived action diagnosticity was higher in the reject condition than the choose condition ($M_{\text{reject}} = 5.88$, 95% CI = [5.67, 6.09] vs. $M_{\text{choose}} = 5.53$, 95% CI = [5.33, 5.74], $t(629) = 2.30$, $p = .022$, $d = 0.19$, 95% CI = [0.03, 0.34]), consistent with H2. Action diagnosticity positively predicted the size of the preference gap ($b_{\text{diagnosticity}} = 0.43$, $SE = 0.04$, $p < .001$) and reduced the coefficient of framing on the preference gap when included in the regression model (from $b = 0.46$, $p = .008$ to $b = 0.31$, $p = .052$). Action diagnosticity mediated the framing effect (indirect effect = -0.15, 95% CI = [-0.29, -0.03]).

Experienced dissonance ($\alpha = .90$). Different from Study 2a, participants in the reject condition reported more dissonance than participants in the choose condition ($M_{\text{reject}} = 2.78$, 95% CI = [2.59, 2.97] vs. $M_{\text{choose}} = 2.35$, 95% CI = [2.16, 2.54]), $t(629) = 3.14$, $p = .002$, $d = 0.25$, 95% CI = [0.09, 0.41]). However, more dissonance predicted a *smaller* post-decision preference gap ($b_{\text{dissonance}} = -0.40$, $SE = 0.05$, $p < .001$), and including dissonance in the regression model *increased* the coefficient of framing on the preference gap (from $b = 0.46$, $p = .008$ to $b = 0.63$, $p < .001$). In other words, in this study, the framing effect was found *despite* participants reporting stronger dissonance in the reject condition than in the choose condition.

Decision closure. Decision closure did not moderate the framing effect (interaction $F(1, 627) = 0.30$, $p = .59$) in this study.

Discussion

We replicated the framing effect in Studies 2a and 2b with different stimuli, further supporting H1. Additionally, both studies provide process evidence for the role of action diagnosticity underlying the framing effect, in line with H2 and the inference account. Meanwhile, the dissonance measures did not accord with the framing effect.

Specifically, while experienced dissonance differed significantly between choose and reject conditions in both studies, they differed in opposite directions in the two studies, yet did not contribute to the framing effect in either study. In Study 2b, in particular, even when dissonance differed in the same direction as would be expected by the dissonance account, the framing effect occurred *despite* dissonance differences. In other words, even if framing affects dissonance in some ways, the dissonance account cannot explain the current framing effect. Last, it should be noted that average dissonance was low ($M_s < 2.80$) in all conditions, further casting doubt on the relevance of dissonance to the framing effect. By contrast, action diagnosticity was moderately high ($M_s > 5.50$) in all conditions, while both were measured on 1-10 scales.

Last, on the role of perceived decision closure, our results in these two studies were inconclusive. Decision closure amplified the framing effect in Study 2a but not in Study 2b.

When we included all the above process variables in several additional replications reported in the SOM, we found similar patterns of results to those reported above. These include consistent mediations by action diagnosticity, no consistent results from experienced dissonance, and a directional moderating role of perceived decision closure, which will be summarized and further discussed in the GD.

The analyses in Studies 2a and 2b provide useful insights that help adjudicate the two major accounts' relative fit with the current framing effect. However, these analyses do not establish causal evidence (Fieldler, Schott, & Meiser, 2011; Fieldler, Harris, & Schott, 2018).

In the subsequent studies, we turn to testing causal evidence by manipulating theoretical moderators (Spencer et al., 2005) of the inference account.

Study 3—Option Similarity

In Study 3, we manipulated attribute similarity between options to test a key boundary condition of the framing effect (H3). We pre-registered the study at <https://aspredicted.org/blind.php?x=zi5jz3>.

Method

Participants. We recruited 533 undergraduate students from a large public university for course credit and obtained 470 valid responses ($M_{\text{age}} = 21$, 167 male, 302 female, 1 non-binary/others) after standard screening.

Procedure. Participants were randomly assigned to one of 2 (framing: choose vs. reject) \times 2 (option similarity: moderate vs. high) between-subjects conditions. Each participant was presented with two donburi bowls (*i.e.*, rice topped with protein and vegetables) as lunch options on a restaurant’s ordering page. We selected two pairs of donburi bowls, a dish familiar to this population. Each bowl contained seven ingredients. The moderately similar pair differed in four ingredients; the highly similar pair differed in one ingredient only (Table 5). In the choose conditions, participants were asked to drag the preferred option into a box labeled “cart.” In the reject conditions, participants were asked to drag the less preferred option into a box labeled “bin.” After the decision, participants were asked to indicate how much they liked each option on a 9-point scale from 1 (*very little*) to 9 (*very much*).

Table 5. Listed ingredients in the donburi bowls.

Moderately Similar Options:	<i>rice, teriyaki chicken, egg, sweet corn, mushroom, lettuce, sesame</i>
------------------------------------	---

	<i>rice, <u>grilled beef sirloin</u>, egg, <u>edamame</u>, <u>pickles</u>, <u>cucumber</u>, sesame</i>
Highly Similar Options:	<i>rice, teriyaki chicken, egg, sweet corn, mushroom, <u>lettuce</u>, sesame</i>
	<i>rice, teriyaki chicken, egg, sweet corn, mushroom, <u>cucumber</u>, sesame</i>

Manipulation check and post-test. After rating both bowls, participants were asked to rate how different the options were on a 10-point scale from 1 (*very similar*) to 10 (*very different*). The results confirm our intended manipulation, showing that the options in the moderate-similarity conditions were deemed more different, $M_{\text{moderately-similar-pair}} = 5.82$, 95% CI = [3.45, 4.04] vs. $M_{\text{highly-similar-pair}} = 3.75$, 95% CI = [5.53, 6.12], $t(468) = 9.71$, $p < .001$, $d = 0.89$, 95% CI = [0.70, 1.08]. Additionally, a post-test ($N = 253$) found that the two bowls in the high-similarity condition were also evaluated as more similar in attractiveness (on a 7-point scale from 1 (*not similar*) to 7 (*very similar*)) than were those in the moderate-similarity condition, $M_{\text{moderately-similar-pair}} = 4.64$, 95% CI = [4.34, 4.93] vs. $M_{\text{highly-similar-pair}} = 5.48$, 95% CI = [5.18, 5.77], $t(251) = 3.96$, $p < .001$, $d = 0.50$, 95% CI = [0.25, 0.75].

Exploratory measures. We included three sets of exploratory measures. These measures were conceptually consistent with those in Studies 2a and 2b but were measured at the general decision level without specifying the related actions and options. Since our main goal in this study was to test a boundary condition, we made the wording changes to shorten the study, reducing four questions for each variable down to one (see full detail in Table S1.3, SOM). However, acknowledging that these changes might reduce the measures' sensitivity, we treated and pre-registered these shortened measures as exploratory.

Results

As predicted, high option similarity mitigated the framing effect in a 2-way ANOVA, $F(1, 466) = 4.95$, $p = .027$, $\eta_p^2 = .01$. The framing effect was replicated in the moderate-similarity conditions, $t(231) = 2.75$, $p = .006$, $d = .36$, 95% CI = [0.10, 0.62], and attenuated

in the high-similarity conditions, $t(235) = 0.06, p = .95$ (Figure 3; Tables 1 & 6).

Unsurprisingly, we also found a main effect of framing ($\Delta_{\text{reject}} = 1.89, 95\% \text{ CI} = [1.69, 2.08]$ vs. $\Delta_{\text{choose}} = 1.62, 95\% \text{ CI} = [1.42, 1.81]$), $F(1, 466) = 4.60, p = .033, \eta_p^2 = .01$, and a main effect of manipulated option similarity, with a generally larger overall preference gap in the moderate-similarity conditions than in the high-similarity conditions ($\Delta_{\text{moderate-similarity-pair}} = 2.03, 95\% \text{ CI} = [1.84, 2.22]$ vs. $\Delta_{\text{high-similarity-pair}} = 1.48, 95\% \text{ CI} = [1.29, 1.67]$), $F(1, 466) = 16.78, p < .001, \eta_p^2 = .04$.

DIP screening excluded similar ratios of participants in the two frames (2% vs. 1%), $\chi^2(1, N = 470) = 0.16, p = .69$, without significantly affecting the results. Again, high option similarity mitigated the framing effect in a 2-way ANOVA, $F(1, 459) = 4.03, p = .045, \eta_p^2 = .01$; the framing effect was replicated in the moderate-similarity conditions, $t(225) = 2.59, p = .010, d = 0.35, 95\% \text{ CI} = [0.09, 0.61]$, and attenuated in the high-similarity conditions, $t(234) = 0.06, p = .96$ (Figure 3; Tables 1 & 6).

Figure 3. The framing effect was attenuated in a decision between highly similar options. All error bars represent 95% confidence intervals.

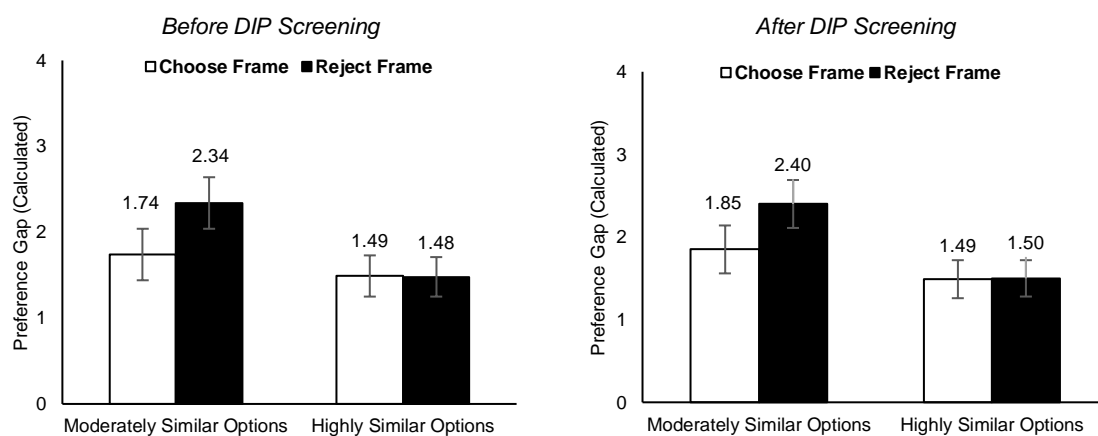


Table 6. Preference ratings by option and by frame in Study 3.

		<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
		Choose Frame	Reject Frame	Choose Frame	Reject Frame
Moderately similar options	Wanted Option	7.12 [6.90, 7.33]	7.31 [7.09, 7.53]	7.16 [6.95, 7.38]	7.32 [7.10, 7.55]
	Unwanted Option	5.38 [5.05, 5.70]	4.97 [4.64, 5.31]	5.31 [4.99, 5.64]	4.93 [4.60, 5.26]
Highly similar options	Wanted Option	6.78 [6.55, 7.01]	7.15 [6.93, 7.37]	6.78 [6.55, 7.01]	7.15 [6.93, 7.37]
	Unwanted Option	5.29 [4.99, 5.60]	5.67 [5.38, 5.96]	5.29 [4.99, 5.59]	5.65 [5.36, 5.94]

Does this moderation merely reflect a “floor effect” in the high-similarity conditions?

That was not the case because the preference gaps were calculated as difference scores from the preference ratings (in Table 6)—rather than measured on an artificially truncated scale. In fact, as Table 6 shows, the smaller preference gaps in the high-similarity conditions resulted from closer ratings between options, all of which fell in the upper middle range of the 9-point (1-9) scale. Further analyses confirmed that while most participants assigned different ratings to the two options in both conditions, more participants in the high-similarity conditions than in the moderate-similarity conditions assigned the same rating to both options (19% vs. 11%), $\chi^2(1, N = 470) = 6.32, p = .012$, consistent with the post-test results.

The three sets of exploratory process measures did not produce statistically meaningful results. In fact, we found across multiple supplemental studies that measuring these variables at the decision level without specifying the concrete actions and options was generally ineffective (section 3, SOM; Tables S3.1 and S3.2). These results and their implications will be further discussed in the GD.

Discussion

The framing effect diminished in a decision between options with highly overlapping attributes, supporting H3 and further corroborating the inference mechanism.

These results are not the most compatible with the dissonance account. Dissonance theory has stressed an “equal attractiveness” antecedent for post-decision dissonance. This antecedent was not satisfied in the moderate-similarity conditions, in which the two options were significantly dissimilar in attractiveness (as shown in the post-test), yet the framing effect still occurred in those conditions. These results further challenge the fit between the dissonance account and the effect of framing on preference modulation. We also conducted an further analyses assessing the role of the “equal attractiveness” antecedent on the framing effect across all our studies (see Figure 7), which will be reported and discussed in the GD.

Studies 4a and 4b—The Valence Moderator

Next, in line with H4, we examined choice-set valence as another moderator of the framing effect. The framing effect has emerged in decisions involving primarily attractive options, and we expected it to be mitigated or even reversed in decision involving unattractive options. We tested these predictions using images in Study 4a and words in Study 4b. Both used a factorial design with 2 (framing: choose vs. reject) \times 2 (choice-set valence: positive vs. negative) between-subjects conditions. The studies were pre-registered at https://aspredicted.org/HS4_NZZ and https://aspredicted.org/DZ5_6BS , respectively.

Study 4a

Method

Participants. We recruited 1,400 Prolific participants, received 1,441 entries, and obtained 1,360 valid responses ($M_{\text{age}} = 40$, 640 male, 696 female, 24 non-binary/others) after standard screening.

Procedure. We selected four photographs from the Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017), an open-access stimulus set that provides 900 color images with validated valence and arousal ratings. Table 7 displays the four photographs, all

selected under the “Scene” category and with similar arousal ratings (between 4.23 and 4.43 out of 7). For their valence (positive vs. negative) ratings, the pair in the positive conditions were rated 6.12 and 6.16 and the pair in the negative conditions were rated 2.24 and 2.29 (out of 7; all images and ratings are available at <https://db.tt/yYTZYCga>; Kurdi et al., 2017).

Table 7. Study 4a used four photographs from the Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017).

Positive Images		Negative Images	
			

Participants were randomly assigned to one of four conditions. Each participant was presented with a pair of photographs, both positive or both negative. In the choose condition, participants were asked to click on their preferred photograph. In the reject condition, participants were asked to click on their less preferred photograph. After making the decision, participants were asked to rate each photograph on a bipolar scale from -10 (*Dislike this photograph very much*) to 10 (*Like this photograph very much*). No exploratory measures were included.

Results

First, before DIP screening, we found the predicted framing \times valence interaction, $F(1, 1356) = 8.06, p = .005, \eta_p^2 = .01$. The framing effect was replicated in the positive-images conditions, $t(675) = 4.20, p < .001, d = 0.32, 95\% \text{ CI} = [0.17, 0.48]$, and diminished in the negative-images conditions, $t(681) = -0.25, p = .80$ (Table 1 & 8).

The DIP screener resulted in similar attritions in the choose and reject frames (2% vs. 3%), $\chi^2(1, N = 1360) = 0.59, p = .44$. The results held after DIP screening, including the framing x valence interaction, $F(1, 1320) = 6.08, p = .014, \eta_p^2 = .01$, the framing effect in the positive-images conditions, $t(660) = 4.50, p < .001, d = 0.35, 95\% \text{ CI} = [0.20, 0.50]$, and the diminished effect in the negative-images conditions, $t(660) = 0.62, p = .54$ (Table 1 & 8).

Figure 4. The framing effect diminished when both images were less desirable. All error bars represent 95% confidence intervals.

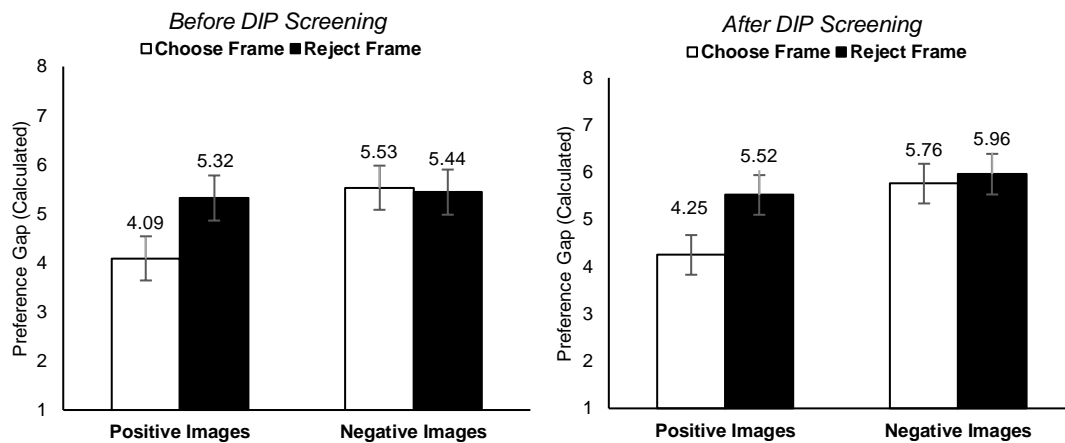


Table 8. Preference ratings by option and by frame in Study 4a.

		<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
		Choose Frame	Reject Frame	Choose Frame	Reject Frame
Positive Images	Wanted Option	6.70 [6.31, 7.09]	7.29 [6.90, 7.69]	6.74 [6.35, 7.13]	7.41 [7.01, 7.81]
	Unwanted Option	2.61 [2.21, 3.02]	1.98 [1.57, 2.39]	2.50 [2.11, 2.88]	1.90 [1.50, 2.29]
Negative Images	Wanted Option	-1.15 [-1.54, -0.75]	-0.70 [-1.09, -0.30]	-1.06 [-1.45, -0.67]	-0.62 [-1.02, -0.22]
	Unwanted Option	-6.67 [-7.08, -6.27]	-6.14 [-6.55, -5.73]	-6.81 [-7.20, -6.43]	-6.58 [-6.97, -6.19]

Study 4b

Method

Participants. We recruited 1,600 Prolific participants, received 1,651 entries, and obtained 1,571 valid responses ($M_{\text{age}} = 39$, 858 male, 688 female, 25 non-binary/others) after standard screening.

Procedure. We used 10 pairs of words (Table 9). All were taken from Perfecto et al. (2017), who matched up pairs of “extremely negative,” “slightly negative,” “slightly positive,” and “extremely positive” words based on valence ratings initially validated in Bellezza et al. (1986). We only used the extreme pairs (and not the “slightly positive” and “slightly negative” pairs), expecting that the more extremely evaluated stimuli may facilitate the detection of a stronger moderation—with the framing effect potentially reversed in the “extremely negative” conditions.

Table 9. Word pairs in Study 4b, equated in valence ratings by Perfecto et al. (2017).

Extremely Positive Words	Extremely Negative Words
Joy vs. Kiss	Murderer vs. Tumor
Pleasure vs. Vacation	Poison vs. Slaughter
Family vs. Laughter	War vs. Maggot
Paradise vs. Sunrise	Cancer vs. Funeral
Romantic vs. Love	Lice vs. Suicide

Participants were randomly assigned to one of the four conditions. Each participant was asked to make five decisions, each involving a pair of words. The five pairs were sequentially presented in randomized order, one pair per page. In the choose condition, participants were asked to mark their preferred word within each pair with a green tick “✓”. In the reject condition, participants were asked to mark their less preferred word within each pair with a red cross “×”. After the decision, participants were asked to rate each word on a bipolar scale similar to that in Study 4a. Each participant’s post-decision preference gap was averaged from ratings of all five pairs of words. No exploratory measures were included.

Results

First, before DIP screening, we found the predicted framing \times valence interaction, $F(1, 1567) = 44.93, p < .001, \eta_p^2 = .03$. The framing effect was again replicated in the positive-words conditions, $t(784) = 7.39, p < .001, d = 0.52, 95\% \text{ CI} = [0.39, 0.67]$, and reversed in direction in the negative-words conditions, $t(783) = -1.66, p = .097, d = -0.12, 95\% \text{ CI} = [-0.26, 0.02]$ (Figure 5; Table 1 & 10).

Since each participant made five decisions, the DIP screener excludes *any* participant with decision-preference inconsistency in *any* word pair. This led to higher attrition ratios in both frames (11% vs. 9%), $\chi^2(1, N = 1571) = 2.30, p = .13$ compared with earlier studies. The results held after DIP screening, including the framing \times valence interaction, $F(1, 1409) = 56.12, p < .001, \eta_p^2 = .04$, the framing effect in the positive-words conditions, $t(724) = 7.20, p < .001, d = 0.54, 95\% \text{ CI} = [0.39, 0.69]$, and the reversed effect in the negative-words conditions, $t(685) = -3.11, p = .002, d = -0.24, 95\% \text{ CI} = [-0.39, -0.09]$ (Figure 5; Table 1 & 10).

Figure 5. The framing effect was reversed when both words were negative in Study 4b. All error bars represent 95% confidence intervals.

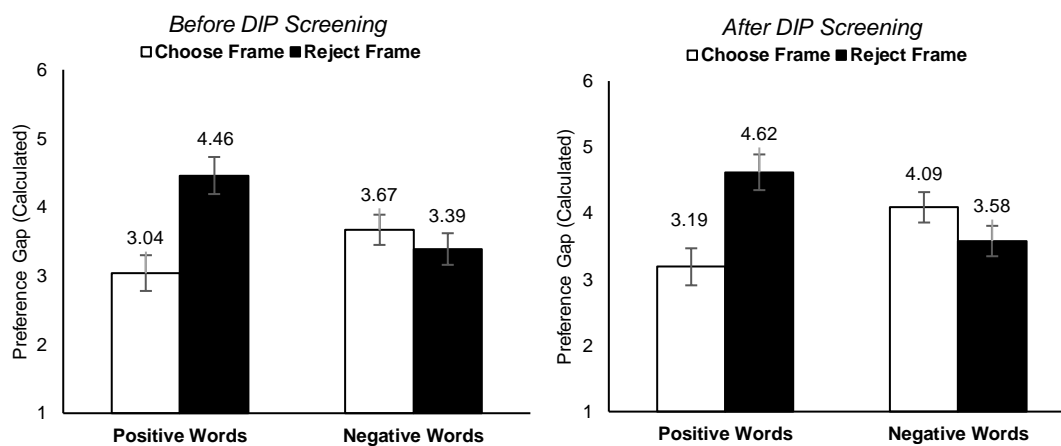


Table 10. Preference ratings by option and by frame in Study 4b.

		<i>Before DIP Screening</i>		<i>After DIP Screening</i>	
		Choose Frame	Reject Frame	Choose Frame	Reject Frame
Positive Words	Wanted Option	7.63 [7.40, 7.86]	7.16 [6.93, 7.40]	7.80 [7.57, 8.03]	7.29 [7.05, 7.52]
	Unwanted Option	4.60 [4.25, 4.94]	2.70 [2.36, 3.05]	4.61 [4.25, 4.96]	2.67 [2.30, 3.03]
Negative Words	Wanted Option	-3.52 [-3.88, -3.18]	-3.55 [-3.90, -3.20]	-3.39 [-3.76, -3.03]	-3.62 [-3.97, -3.26]
	Unwanted Option	-7.19 [-7.47, -6.92]	-6.94 [-7.22, -6.67]	-7.48 [-7.75, -7.22]	-7.20 [-7.46, -6.94]

Discussion

Similar results were found in both Study 4a, using images, and Study 4b, using words. While the original framing effect was consistently replicated in the positive conditions, it no longer emerged in the negative conditions—when decisions involving primarily unattractive options. These results support the moderating role of choice-set valence to the effect, in line with H4. Additionally, in an earlier, less powered version of Study 4b employing the same word pairs and a different cover story (in which the words were introduced as temporary hand tattoo options), we found principally similar results to those in Study 4b (see Study S2 in the SOM). Together, these results indicate that the valence moderator is robust to idiosyncratic differences in stimuli and contexts.

The effect sizes in Studies 4a and 4b provide additional insights. As can be seen in Tables 8 and 10, the positive and negative stimuli in Study 4b received more extreme ratings than those in Study 4a. Correspondingly, both the framing effect in the positive conditions ($d = .52$) and its directionally reversed effect in the negative conditions ($d = -0.12$) in Study 4b were more sizable than those in Study 4a ($d = .32$ and $d = -0.02$). In fact, in the negative conditions, the original framing effect was only attenuated in Study 4a, whereas it was reversed in direction in Study 4b. These patterns coalesce with previous framing studies (e.g.,

on post-decision confidence, Study 2 in Perfecto et al. 2017), in which the reversal of a framing effect was facilitated by stimuli of more extreme valence.

We have now obtained empirical evidence for the four hypotheses derived from the inference account. Although dissonance was not measured in Studies 4a and 4b, it would be improbable for the dissonance account to provide a coherent explanation for the current results. This is because our earlier results from Studies 2a and 2b already revealed that the simple framing effect in positive-valence conditions was incompatible with the dissonance account, let alone its reversal in negative-valence conditions.

Study 5—An Alternative Design to Assess Potential Confounds

Thus far, we have replicated the framing effect in six studies using true randomization, large samples, and diverse stimuli. These results indicate that the effect was robust to random errors (e.g., participant' idiosyncratic differences in preference volatility or inattentiveness). Additionally, the framing effect remained mostly unchanged before and after DIP screening in all six studies, indicating that the effect was not significantly affected by systematic errors either (e.g., inflated gap calculations due to unequal distributions of preference reversals).

In Study 5, we sought to hone in on the role of systematic error on the framing effect using yet another method. This method was inspired by one of the tested solutions offered by Chen and Risen (2010) to address previous confounds in the FCP (Brehm, 1956). We modified their solution to fit our framing paradigm. Specifically, we added a new set of control conditions, in which the order between the decision and preference ratings reversed such that the decision *cannot* possibly influence the preference ratings in these conditions. If the framing effect primarily reflects different degrees of decision-induced preference

modulation, as we theorized, then it should not emerge in these control conditions. In contrast, if framing systematically affected the extent to which a participant's decision would misrepresent their existing preferences, then the effect should emerge similarly in both experimental and control conditions. We pre-registered this study and its full analysis plan at https://aspredicted.org/TK2_7Q6.

Method

Participants. We recruited 1,400 Prolific participants, received 1,484 entries, and obtained 1,432 valid responses ($M_{\text{age}} = 38$, 623 male, 774 female, 35 non-binary/others) after standard screening.

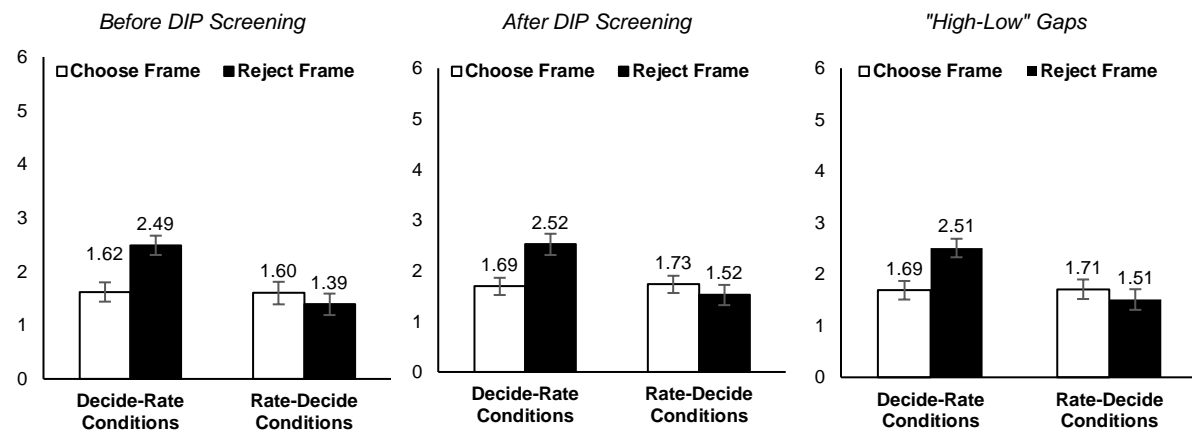
Procedure. Participants were randomly assigned to 2 (framing: choose vs. reject) \times 2 (order: decide-rate vs rate-decide) between-subjects conditions. Each participant was presented with a pair of words, “*pleasure*” and “*vacation*,” taken from Study 4b. The decide-rate (D-R) conditions were similar to the positive-valence conditions in Study 4b. The rate-decide (R-D) conditions mirrored the D-R conditions. The only difference was that participants were first asked to rate the two words, and then to make a decision. In other words, in the R-D conditions, the decision followed preference ratings, rather than preceding them. Decision-induced preference modulation could only occur in the D-R conditions. However, if framing influences the likelihood of preference reversal, then it should impact both D-R and R-D conditions.

Results

Following all three sets of pre-registered analyses, the framing effect was observed only in the D-R conditions and not in the R-D conditions. Below, we report them in order.

First, before DIP screening, the framing effect emerged in the D-R conditions, $t(714) = 6.68, p < .001, d = 0.50, 95\% \text{ CI} = [0.35, 0.65]$, and not in the R-D conditions, $t(714) = -1.46, p = .15$. The mitigation of the framing effect in the R-D conditions was validated by an interaction between framing and (D-R vs. R-D) order, $F(1,1428) = 30.20, p < .001, \eta_p^2 = .02$ (Figure 6 left panel; Tables 1 & 11). Moreover, the framing effect was primarily driven by gap differences within the reject conditions (simple effect between the black bars: $t(718) = 7.68, p < .001, d = 0.57, 95\% \text{ CI} = [0.42, 0.72]$). The two gaps in the choose conditions did not significantly differ (simple effect between the white bars: $t(710) = 0.17, p = .87$).

Figure 6. Post-decision preference gaps in Study 5. Each panel corresponds to one set of pre-registered analysis. The framing effect was expected in the D-R conditions and never expected in the R-D conditions. Positive preference gaps (bars above 0) were expected in all conditions irrespective of framing or the order between decision and ratings.



Second, DIP screening excluded similar ratios of participants in choose and reject conditions (3% vs. 3%), $\chi^2(1, N = 1432) = 0.26, p = .61$. The results remained similar including the framing effect in the D-R conditions, $t(703) = 6.52, p < .001, d = 0.49, 95\% \text{ CI} = [0.34, 0.64]$, its absence in the R-D conditions, $t(684) = -1.44, p = .15$ (Tables 1 & 11), and

the interaction between framing and order, $F(1,1387) = 28.93, p < .001, \eta_p^2 = .02$ (Figure 6, middle panel).

Third, we computed an alternative “high-low gap” by subtracting the lower rating from the higher rating between options without DIP screening. In this calculation, the minuend and subtrahend were solely determined by preference ratings, independent of the decision outcome (which determined the “wanted” and “unwanted” options in all previous calculations of preference gaps). This “high-low gap” reflects the face-value preference gap between options irrespective of the decision outcome, since the minuend was always the higher rating. This means that each bar in the R-D conditions represents a baseline preference gap plus random error, whereas each bar in the D-R conditions represents a preference gap caused by having made a (choose vs. reject) decision, on top of the baseline preference gap plus random error. Again, all our results remained, including the framing effect in the D-R conditions, $t(714) = 6.61, p < .001, d = 0.49, 95\% \text{ CI} = [0.34, 0.64]$, its absence in the R-D conditions, $t(714) = -1.46, p = .15, d = -0.11, 95\% \text{ CI} = [-0.25, 0.04]$ (Tables 1 & 11), and the significant interaction between framing and decision-rating order, $F(1,1428) = 30.13, p < .001, \eta_p^2 = .02$ (Figure 6, right panel).

Table 11. Preference ratings by option and by frame in Study 5.

		<i>Before DIP Screening</i>		<i>After DIP Screening</i>		<i>“High-low Gaps”</i>	
		Choose Frame	Reject Frame	Choose Frame	Reject Frame	Choose Frame	Reject Frame
Decide-Rate	Wanted Option	8.47 [8.33, 8.61]	8.91 [8.78, 9.05]	8.51 [8.37, 8.65]	8.92 [8.79, 9.06]	Higher-Rated Option	8.51 [8.37, 8.64] 8.91 [8.78, 9.05]
	Unwanted Option	6.84 [6.63, 7.06]	6.42 [6.21, 6.63]	6.82 [6.60, 7.03]	6.40 [6.19, 6.62]	Lower-Rated Option	6.83 [6.62, 7.04] 6.41 [6.19, 6.62]
Rate-Decide	Wanted Option	8.97 [8.83, 9.11]	8.99 [8.86, 9.13]	9.06 [8.92, 9.19]	9.06 [8.93, 9.20]	Higher-Rated Option	9.02 [8.89, 9.15] 9.06 [8.93, 9.19]

Unwanted Option	7.37 [7.13, 7.61]	7.61 [7.37, 7.84]	7.33 [7.09, 7.58]	7.55 [7.31, 7.79]	Lower-Rated Option	7.29 [7.06, 7.53]	7.55 [7.31, 7.78]
--------------------	----------------------	----------------------	----------------------	----------------------	-----------------------	----------------------	----------------------

Discussion

Study 5 confirms that the framing effect primarily reflects more preference changes *caused* by the reject (vs. choose) decision instead of potential calculation confounds associated with unequal distributions of preference reversals between frames. These results corroborate our previous results using the DIP screening method. Besides Study 5, we also report Study S3 in the SOM, which used the same design as Study 5, the same smartphone cases in Study 1, and yielded principally similar results; although it was under-powered to detect a significant 2-way interaction between framing and order, it replicated both the framing effect in the R-D conditions and the null effect in the D-R conditions.

General Discussion

Across seven pre-registered experiments, we reported that framing a decision as a rejection intensified post-decision preferences compared to framing the same decision as a choice. This framing effect was observed across diverse positive stimuli including products, art prints, food, photos, and words. The larger preference gap in the reject (vs. choose) frame was consistent with the greater perceived diagnosticity of the reject (vs. choose) action. Moreover, the framing effect was moderated by two theoretical moderators, option similarity and choice-set valence. Together, our findings support our hypotheses derived from the inference account, and by necessity, support the inference account for preference modulation. In contrast, the dissonance account received little support when tested in these studies.

The framing effect was further replicated in another six pre-registered supplemental studies ($N = 3,679$) reported in the SOM: Studies S1, S2, and S3 were similar to Studies 1,

4b, and 5, respectively, with minor procedural differences. Studies S4, S5, and S6 extended the effect's robustness to larger choice sets in which two out of four art prints were chosen or rejected (S4), and to procedures that required participants to write down their preferences before the decision (S5 & S6). Their main results are summarized in Table S2.1 (SOM). Across all 13 studies, we did not find the size of the framing effect to depend on specific procedural operationalizations of the framing manipulation (e.g., dragging, clicking, or marking options; summarized in Table S1.2, SOM).

Other Decision Consequences

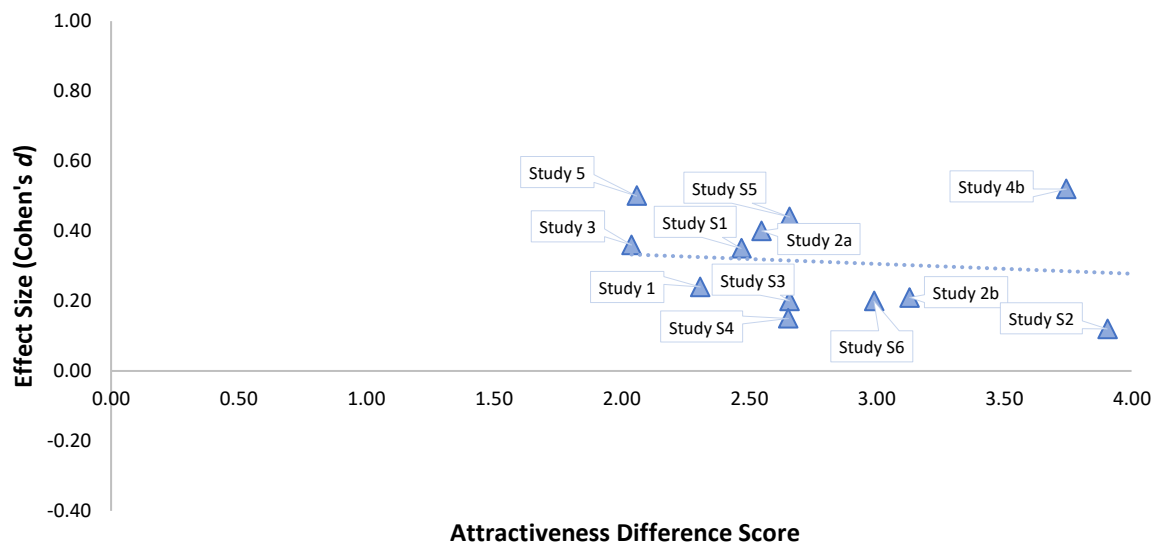
Next, we report findings of second-order theoretical interest. First, as briefly mentioned after Study 1, we examined (as pre-registered) whether the framing effect was driven more by **the wanted option or the unwanted option**. Across all thirteen studies, we found that both options contributed to the preference modulation, and neither option had a consistently larger contribution (Table S4.1, SOM).

Second, we examined whether framing affected **choice shares** (i.e., the binary decision outcome) and **average preference ratings**. We did not find the choice shares of options to differ systematically between the choose and reject conditions (Table S4.3, SOM). Nor did the post-decision preference ratings systematically differ between choose and reject conditions (Table S4.2, SOM). Therefore, the current framing effect was not contingent on framing altering the decision outcomes or shifting the average evaluation.

Third, returning to the “equal attractiveness” antecedent proposed by dissonance theory, we explored whether **the similarity in attractiveness between options** affected the size of the framing effect. Since we did not equate option attractiveness within each pair of options in most studies, their differences varied considerably. We thus computed an

“attractiveness difference score” in each study based on the average preference ratings of each option (see Table S4.4, SOM). Then, we plotted the score in each study against the observed size of the framing effect in Figure 7. If “equal attractiveness” was indeed an antecedent for preference modulation, and hence a necessary condition to observing the framing effect, then observed effect sizes should cluster on the far left of the x -axis where the options are equal in attractiveness. Moreover, average effect size should quickly decline as the options diverge in attractiveness, with a sharp downward slope representing a negative correlation between attractiveness difference scores and effect sizes.

Figure 7. In the test conditions of all thirteen studies, observed effect sizes did not correlate with differences in attractiveness between options ($r = -0.165$, $p = .59$).



However, as Figure 7 shows, the framing effect occurred in a wide range of attractiveness difference scores, mostly away from zero. Their effect sizes did not correlate with the attractiveness difference score, Pearson's $r = -0.165$, $p = .59$. These results suggest

that preference modulation did not depend on the “equal attractiveness” antecedent, consistent with the results in Study 3. The results are highly similar when we computed attractiveness differences from choice shares.

What do these results suggest for dissonance theory? There are two interpretations. One, let’s assume that dissonance theorists were right about the “equal attractiveness” antecedent (again, even although prior studies did not adequately test this assumption; see Footnote 3). Then, perhaps the options in our studies do not meet the stringent criteria of being “equal” in attractiveness while sufficiently “dissimilar” in attributes to triggers dissonance—a rather narrow range of stimuli in real-world settings. This suggests that dissonance is not as broadly applicable as has been presumed. Alternatively, perhaps the “equally attractiveness” antecedent was a faulty, unnecessary assumption for dissonance theory. In that case, the results reported in Study 3 and Figure 7 are no longer evidence against the dissonance account, even though the dissonance account remains incompatible with the framing effect based on results in Studies 2a & 2b (and more supplemental studies finding null results of dissonance, as summarized below). Either way, the dissonance account cannot explain the current framing effect. Dissonance researchers may reconsider and/or re-examine the necessary conditions of dissonance and its boundaries in light of our findings.

Additional Results on Process Variables

In addition to Studies 2a and 2b, we measured perceived action diagnosticity, experienced dissonance, and perceived decision closure in five of the six supplemental studies. Analyses on these measures revealed consistent patterns, which we summarize here using an internal meta-analysis (IMA) conducted on the test conditions of all eight studies in which these variables were measured ($N = 5,852$). While we report all our studies in the

paper and the SOM, we acknowledge and agree with important critiques that the IMA can inadvertently inflate effect sizes (see Vosgerau et al., 2019). Thus, we intend to use the IMAs to succinctly highlight consistent patterns in our findings, not to claim specific effect sizes.

First, we found that **perceived action diagnosticity** was significantly higher in the reject frame than the choose frame in the test conditions of every study when diagnosticity was measured at the action and option level (i.e., in reference to the concrete actions and corresponding options; in Studies 2a, 2b, S1, S5, & S6; Tables S1.3 & S3.1, SOM), whereas null results were found when this variable was measured at the decision level (in Studies 3, S2, & S4; Table S3.1, SOM). We surmise that the generic wording in the decision-level measurement failed to engage participant attention effectively. Overall, perceived diagnosticity differed significantly between framing conditions across all eight studies (average diagnosticity difference = -0.26, $SE = 0.08$, 95% CI = [-0.42, -0.11], $Z = -3.34$, $p < .001$), and mediated the framing effect (indirect effect = -0.10, $SE = 0.04$, 95% CI = [-0.17, -0.03], $Z = -2.70$, $p = .007$), with sizable heterogeneity across studies, $I^2 = 57\%$, 95% CI = [10%, 93%].

Second, **experienced dissonance** did not consistently differ between frames regardless of measurement type (Table S3.2, SOM), nor did it mediate the framing effect (indirect effect = -0.01, $SE = 0.05$, 95% CI = [-0.09, 0.08]), $Z = -0.14$, $p = .89$. These results corroborate our earlier results to show that dissonance was unlikely the primary mechanism for the framing effect. Given that the self-report measures of dissonance produced statistically significant yet contradictory results (rather than null effects) between Study 2a and Study 2b, measurement sensitivity issues alone cannot explain these discrepant results. If alternative measures of dissonance (e.g., physiological arousal; Croyle & Cooper 1983) were

to be used in future studies, researchers should address not only potential measurement sensitivity issues but also the discrepancies in the results across studies.

Third, we found that perceived **decision closure** amplified the framing effect in four out of the eight studies in which it was measured (Studies 2a, S1, S2, & S4; section 3.3, SOM). Overall, the framing effect was stronger at higher levels of decision closure (conditional effect at $M_{\text{closure}} + 1\text{SD} = -0.76$, $SE = 0.16$, 95% CI = [-1.08, -0.45], $Z = -4.75$, $p < .001$) and weaker at lower levels of decision closure (conditional effect at $M_{\text{closure}} - 1\text{SD} = -0.27$, $SE = 0.20$, 95% CI = [-0.65, 0.12], $Z = -1.37$, $p = .17$). Taken together, this moderation pattern suggests that preference updating more readily occurs when a decision is deemed complete. These additional findings are consistent with the inference account. They do not help adjudicate the applicability of the dissonance account because perceived closure may increase or decrease dissonance (e.g., Carmon et al., 2003; Stalder, 2010).

Last, we explored **decision time** differences, a measure imbedded on the decision page in Studies 2b, 3, 4, S2, S4, and S5. In all the test conditions of these studies ($N = 3,122$), we found that the reject decision took longer time than the choose decision (average difference = 1.63, $SE = 0.36$, 95% CI = [0.91, 2.34], $Z = 4.47$, $p < .001$), consistent with the greater processing depth in the reject (vs. choose) frame reported in prior research (e.g., Sokolova & Krishna, 2016). However, this difference emerged not only in the test conditions, but also in the boundary conditions (e.g., Studies 3 & S2; Table S3.6), in which framing did not affect preference gaps. Indeed, decision time differences did not mediate the framing effect (indirect effect < 0.001 , $SE = 0.006$, 95% CI = [-0.01, 0.01], $Z = 0.02$, $p = .98$). Based on these results and related critiques on decision time as a rather crude measure of psychological processes (Evans et al., 2015), we surmise that longer decision time may

accompany the reject frame as a concomitant but is not the most reliable process measure, nor a unique explanation for the framing effect.

Theoretical, Methodological, and Practical Implications

This research makes several contributions. First, empirically, we present the first evidence that choose/reject framing affects post-decision preference modulation. This finding extends previous research, which had focused on the impact of framing on decision outcomes (Dhar & Wertenbroch, 2000; Shafir, 1993). We show that framing can influence post-decision preferences beyond decision outcomes—even without affecting decision outcomes *per se*. Moreover, by establishing that reject framing intensifies preferences more than choose framing, we join increasing research on the multifaceted effects of reject framing, such as in eliciting deliberative judgments (Sokolova & Krishna, 2016) and boosting actions' social influence (Nan et al., 2023). The reject frame and its potential implications merit more research in the future.

Second, we address a long-standing theoretical ambiguity surrounding the psychological mechanisms of preference modulation. Despite common references to preference modulation as “a dissonance phenomenon” in the literature, the dissonance account lacked support when put to the test in the current context, whereas the inference account readily predicted and explained all our results. These findings show that preference modulation can often occur without invoking dissonance. These findings also suggest that preference modulation may be more commonplace than previously conjectured based on the assumed dissonance explanation. To be clear, our results should not be interpreted as refuting dissonance as a plausible explanation for *all* preference modulation or all phenomena of rationalization. Rather, our results suggest that dissonance is probably not the best

explanation for preference changes induced by ordinary daily decisions—such as those examined in our studies. Such quotidian decision contexts are arguably unlikely to elicit from people extraordinary emotional or motivational conflicts after all.

Indeed, many psychologists consider the inference account more parsimonious (see Bem, 1967; Laurin & Jettinghoff, 2020) for it does not necessitate any emotional-motivational dynamics. The inference account is based on the simple premise that spontaneous observation-based inferences may operate similarly in intrapersonal and interpersonal processes. As such, our findings align with converging insights that the dissonance and inference accounts may differ in their primary domains of application. Drawing from our own and prior findings, we surmise that the inference mechanism may be generally more relevant to common ordinary decisions, whereas the dissonance mechanism may be more relevant to more extraordinary situations (or, “flagrant behavioral situations,” Simon & Holyoak, 2002; see also Abelson, 1983)—in which overtly attitude-incongruent, painstaking, or embarrassing actions do evoke intense emotions and motivational conflicts.

Another smaller theoretical contribution may stem from our findings on the self-diagnosticity of a decision, which was bolstered by the reject frame. These findings cast a spotlight on the underexplored role of self-diagnosticity (Bodner & Prelec, 2002; Dhar & Wertenbroch, 2012) in modulating the relationship between decisions and preferences. Future research may find it fruitful to investigate situational factors that determine the diagnosticity of decisions and actions, with potential implications on cognitive consistency, identity formation, and behavior change.

Third, this research offers a methodological contribution for studying decision-induced preference modulation. Traditionally, researchers relied on the measuring of pre- vs.

post-choice “preference spreading” in FCP and its variants, which involved laborious (and sometimes confounded) protocols. We used a simpler approach to study preference modulation without having to examine “preference spreading.” Our approach combines a framing paradigm with the monitoring of potential calculation confounds through additional screening and/or additional control conditions. Based on appropriately addressing random and systematic errors, we show that such an approach can be used to test for the presence of preference modulation and its underlying mechanisms. In fact, the current framing effect could be observed only if preference modulation occurred AND if framing affected its degree. As such, our results also serve as an independent validation of preference modulation, corroborating existing evidence (reviewed in Enisman et al. 2021).

For future work studying preference modulation and related effects, it is important that they rule out the possibility that framing affects the “noise” in preference expression through decisions. If future studies use our framing paradigm, then the DIP screening method, easy to implement and minimally costly, should be used in all studies when appropriate. If the DIP screening method detects a significant deviation of results, then the costlier and more conclusive R-D vs D-R design should be used to further evaluate the relative contributions of alternative causes. At the same time, it is imperative that researchers also adhere to rigorous practices in other aspects (e.g., sufficient statistical power and sample size) to minimize random error, another major cause for some spurious earlier findings reported in this literature. We hope our novel methods help move researchers past previous methodological challenges and facilitate more theory testing on the phenomenon of preference modulation.

Finally, this research provides rich insights into the uses and consequences of framing interventions in practical domains. In today's political climate, for example, negative advertising strategies frequently replace positive advertising strategies, reframing choices into rejections. Our findings suggest that the impact of such strategies on voters' future political attitudes is a valid concern, and its impact will in part depend on the desirability of the candidates. Between two desirable candidates, a reject-framed campaign or ballot design may exacerbate the polarization of political attitudes among voters, whereas the opposite may be true when both candidates are primarily undesirable. Policymakers should take these potential ramifications into account when adopting framing interventions.

Our findings also suggest means to proactively leverage framing for positive business and social outcomes. Marketers aiming to enhance customers' post-purchase satisfaction and brand loyalty may highlight the "reject" aspects of the shopping journey. Conversely, health practitioners may highlight the "choose" aspects of patients' decisions involving unpleasant treatment options to bolster post-treatment evaluation, potentially reducing patient attrition in long-term treatment plans.

The Information Inference Account: Clarifications, Limitations, and Future Directions

To reiterate, the inference account suggests that people make attributional inferences upon observing an action, which helps them understand the actor's dispositions and motives either when the actor is another individual or oneself. As such, the framing effect rests on general associations between actions and their diagnosticity—reject actions are typically treated as more diagnostic of an actor's preferences than are choose actions. As such, an accurate perception of the decision process is not necessary for the effect. That is, the actual diagnosticity of a decision process and its perceived diagnosticity are separate issues. Further,

our findings do not imply that the information inference process following reject decisions will necessarily engage more cognitive resources than those following choose decisions. The post-decision processes of inference and updating may have distinct psychological characteristics than the decision-making process per se.

On a related note, some readers may be inclined to interpret the process of self-preference inference and updating primarily as a deliberate and controlled process. This impression could arise as “attribution” and “inference” are often discussed as components of a highly effortful and controlled process of causal reasoning. However, we do not presume that self-preference inferences are necessarily effortful and controlled, nor do our findings rely on such assumptions. In fact, we think it is entirely possible that such processes are spontaneous and habitual given how common and potentially adaptive it is (see Cushman 2020).

Third and related, some might expect self-preference inferences to be contingent on knowing uncertainty about one’s existing preferences, a need to justify one’s decision to others, or high identity-relevance of decision stimuli. While such conjectures have been raised in prior self-perception research, we are skeptical about them as well. We agree that these factors can enhance one’s motivation to engage in self-relevant attributional inferences, especially if the inference process primarily reflects conscious and goal-directed efforts of reasoning. However, their relevance would be diminished if the inference process is spontaneous and habitual. Given our agnosticism about the nature of this process, we do not view the above factors as necessary boundary conditions to the framing effect. In our limited explorations of such factors in the supplemental studies and additional IMA analyses, no supportive evidence for them emerged. Nonetheless, more data are required before one can

conclude whether these factors affect the framing effect and preference modulation in general.

Finally, there are interesting questions awaiting future research. For instance, different interpretations exist on the nature of preference modulation within the inference camp: does it reflect a change in actual preferences (per Cushman, 2020), or rather a change in preference beliefs (per Bem, 1972)? Our current studies cannot address these nuanced points. There are also other emerging accounts of preference changes based on sampling of mental representations (Regenwetter et al., 2011) that are beyond the scope of our empirical investigations. Other questions that we wish to examine in follow-up work include how long the framing effect will persist, whether it will extend beyond immediate self-reported preferences, and if it will influence subsequent incentive-compatible behaviors.

Constraints on Generality

Our results were obtained from thirteen studies involving 8,782 North American adults from screened online study platforms and 1,145 Asian undergraduate students from a large public university. Similar results were found across common decision contexts with diverse choice stimuli, based on generic instructions for decisions and preference ratings. We expect these results to generalize to a broad range of populations and situations. Future replications should consider identified boundary conditions for the framing effect such as high attribute similarity and unattractive choice sets. In our studies, preference ratings were measured immediately following decisions. Even though prior studies found similar magnitudes of preference modulation in hypothetical and incentivized studies, sometimes with a significant time lag between measures (e.g., Sharot et al., 2010), it is possible that the

effect size may be affected by changes in these conditions. We have no reason to believe that the results depend on other characteristics of the participants, materials, or context.

Context

The initial impetus of this research was formed when the first author attended the doctoral seminars offered by Jane Risen and Nick Epley at the University of Chicago Booth School of Business. A few years later, when the second author joined the first author to explore novel framing effects, this interest was further developed into the current research. We acknowledge that this paper benefited tremendously from recent advancements in the literature, particularly the methodological and theoretical critiques by Keith Chen and Jane Risen (2010) on previous research on preference modulation, and the expanded framework of rationalization proposed by Fiery Cushman (2020). We hope that our findings, insights, and methods will further inspire and facilitate future research on the important phenomenon of decision-induced preference modulation.

REFERENCES

- Abelson, R. P. (1983). Whatever became of consistency theory? *Personality and Social Psychology Bulletin*, 9(1), 37–54. <https://doi.org/10.1177/0146167283091006>
- Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology*, 46(2), Article 2. <https://doi.org/10.1037/0022-3514.46.2.294>
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology*, 5(4), 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>
- Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, & Computers*, 18(3), Article 3. <https://doi.org/10.3758/BF03204403>
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183–200. <https://doi.org/10.1037/h0024835>
- Bem, D. J. (1972). Self-Perception Theory. In *Advances in Experimental Social Psychology* (Vol. 6, pp. 1–62). Elsevier. [https://doi.org/10.1016/S0065-2601\(08\)60024-6](https://doi.org/10.1016/S0065-2601(08)60024-6)
- Bem, D. J., & McConnell, H. K. (1970). Testing the Self-Perception Explanation of Dissonance Phenomena: On the Salience of Premanipulation Attitudes. *Journal of Personality and Social Psychology*, 14(1), 23–31. <https://doi.org/10.1037/h0020916>
- Bizer, G. Y., Larsen, J. T., & Petty, R. E. (2011). Exploring the Valence-Framing Effect: Negative Framing Enhances Attitude Strength: Valence Framing. *Political Psychology*, 32(1), 59–80. <https://doi.org/10.1111/j.1467-9221.2010.00795.x>

- Bodner, R., & Prelec, D. (2002). Self-signaling and diagnostic utility in everyday decision making. In I. Brocas & J. Carillo (Eds.), *Collected Essays in Psychology and Economics* (p. 22). Oxford University Press.
- Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3), Article 3. <https://doi.org/10.1037/h0041006>
- Brehm, J. W., & Cohen, A. R. (1959). Re-evaluation of choice alternatives as a function of their number and qualitative similarity. *The Journal of Abnormal and Social Psychology*, 58(3), 373.
- Buiten, M. V., & Keren, G. (2009). Speakers' choice of frame in binary choice: Effects of recommendation mode and option attractiveness. *Judgment and Decision Making*, 4(1), 51–63. <https://doi.org/10.1017/S193029750000070X>
- Carmon, Z., Wertenbroch, K., & Zeelenberg, M. (2003). Option Attachment: When Deliberating Makes Choosing Feel like Losing. *Journal of Consumer Research*, 30(1), 15–29. <https://doi.org/10.1086/374701>
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: Revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, 99(4), 573–594. <https://doi.org/10.1037/a0020217>
- Couch, A., & Keniston, K. (1960). Yeasayers and naysayers: Agreeing response set as a personality variable. *The Journal of Abnormal and Social Psychology*, 60(2), Article 2. <https://doi.org/10.1037/h0040372>
- Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, 43, e28. <https://doi.org/10.1017/S0140525X19001730>

Dhar, R., & Sherman, S. J. (1996). The effect of common and unique features in consumer choice. *Journal of Consumer Research*, 23(3), 193–203.

Dhar, R., & Wertenbroch, K. (2000). Consumer Choice between Hedonic and Utilitarian Goods. *Journal of Marketing Research*, 37(1), Article 1.

<https://doi.org/10.1509/jmkr.37.1.60.18718>

Dhar, R., & Wertenbroch, K. (2012). Self-Signaling and the Costs and Benefits of Temptation in Consumer Choice. *Journal of Marketing Research*, 49(1), Article 1.

<https://doi.org/10.1509/jmr.10.0490>

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, 80(2), 312–325.

<https://doi.org/10.1016/j.neuron.2013.09.007>

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, 67(3), 382–394. <https://doi.org/10.1037/0022-3514.67.3.382>

Enisman, M., Shpitzer, H., & Kleiman, T. (2021). Choice changes preferences, not merely reflects them: A meta-analysis of the artifact-free free-choice paradigm. *Journal of Personality and Social Psychology*, 120(1), 16–29.

<https://doi.org/10.1037/pspa0000263>

Evans, A. M., Dillon, K. D., & Rand, D. G. (2015). Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General*, 144(5), 951–966.

<https://doi.org/10.1037/xge0000107>

Fazio, R. H., Zanna, M. P., & Cooper, J. (1977). Dissonance and self-perception: An integrative view of each theory's proper domain of application. *Journal of*

Experimental Social Psychology, 13(5), 464–479. [https://doi.org/10.1016/0022-1031\(77\)90031-2](https://doi.org/10.1016/0022-1031(77)90031-2)

Festinger, L. (1957). *A theory of cognitive dissonance* (pp. xi, 291). Stanford University Press.

Festinger, L. (1962). Cognitive Dissonance. *SCIENTIFIC AMERICAN*, 14.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2), 203–210.
<https://doi.org/10.1037/h0041593>

Ganzach, Y. (1995). Attribute Scatter and Decision Outcome: Judgment versus Choice. *Organizational Behavior and Human Decision Processes*, 62(1), Article 1.
<https://doi.org/10.1006/obhd.1995.1036>

Gerard, H. B. (1967). Choice difficulty, dissonance, and the decision sequence. *Journal of Personality*.

Green, D. (1974). Dissonance and self-perception analysis of “forced compliance”: When two theories make competing predictions. *Journal of Personality and Social Psychology*, 29(6), 819–828. <https://doi.org/10.1037/h0036205>

Greenwald, A. G. (1975). On the inconclusiveness of “crucial” cognitive tests of dissonance versus self-perception theories. *Journal of Experimental Social Psychology*, 11(5), 490–499. [https://doi.org/10.1016/0022-1031\(75\)90051-7](https://doi.org/10.1016/0022-1031(75)90051-7)

Gu, Y., Botti, S., & Faro, D. (2013). Turning the Page: The Impact of Choice Closure on Satisfaction. *Journal of Consumer Research*, 40(2), Article 2.
<https://doi.org/10.1086/670252>

- Gu, Y., Botti, S., & Faro, D. (2018). Seeking and Avoiding Choice Closure to Enhance Outcome Satisfaction. *Journal of Consumer Research*, 45(4), Article 4.
<https://doi.org/10.1093/jcr/ucy025>
- Harmon-Jones, E., & Harmon-Jones, C. (2002). Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes. *Personality and Social Psychology Bulletin*, 28, 711–723.
<https://doi.org/10.1177/0146167202289001>
- Harmon-Jones, E., & Mills, J. (1999). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology*. (pp. 3–21). American Psychological Association. <https://doi.org/10.1037/10318-001>
- Houston, D. A., Sherman, S. J., & Baker, S. M. (1991). Feature matching, unique features, and the dynamics of the choice process: Predecision conflict and postdecision satisfaction. *Journal of Experimental Social Psychology*, 27(5), 411–430.
[https://doi.org/10.1016/0022-1031\(91\)90001-M](https://doi.org/10.1016/0022-1031(91)90001-M)
- Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, 107(51), Article 51.
<https://doi.org/10.1073/pnas.1011879108>
- Izuma, K., & Murayama, K. (2013). Choice-Induced Preference Change in the Free-Choice Paradigm: A Critical Methodological Review. *Frontiers in Psychology*, 4.
<https://doi.org/10.3389/fpsyg.2013.00041>

- Jones, E. E., & Davis, K. E. (1965). From Acts To Dispositions The Attribution Process In Person Perception. In *Advances in Experimental Social Psychology* (Vol. 2, pp. 219–266). Elsevier. [https://doi.org/10.1016/S0065-2601\(08\)60107-0](https://doi.org/10.1016/S0065-2601(08)60107-0)
- Kahneman, D. (2011). *Thinking, fast and slow*.
http://ebook.3m.com/library/neworleanspubliclibrary-document_id-dpt1489
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28(2), 107.
- Kessler, E. H., Ford, C. M., & Bailey, J. R. (1996). Object valence as a moderator of the framing effect on risk preference. *Journal of Economic Behavior & Organization*, 30(2), Article 2. [https://doi.org/10.1016/S0167-2681\(96\)00860-8](https://doi.org/10.1016/S0167-2681(96)00860-8)
- Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, 49, 457–470.
- Laran, J., & Wilcox, K. (2011). Choice, Rejection, and Elaboration on Preference-Inconsistent Alternatives. *Journal of Consumer Research*, 38(2), 229–241.
<https://doi.org/10.1086/659040>
- Laurin, K., & Jettinghoff, W. M. (2020). What kind of rationalization is system justification? *Behavioral and Brain Sciences*, 43, e39. <https://doi.org/10.1017/S0140525X19002243>
- Meloy, M. G., & Russo, J. E. (2004). Binary choice under instructions to select versus reject. *Organizational Behavior and Human Decision Processes*, 93(2), Article 2.
<https://doi.org/10.1016/j.obhdp.2003.12.002>
- Mizerski, R. W. (1982). An Attribution Explanation of the Disproportionate Influence of Unfavorable Information. *Journal of Consumer Research*, 9(3), 301.
<https://doi.org/10.1086/208925>

Nan, L. X., Park, S. K., & Yang, Y. (2023). Rejections Are More Contagious than Choices: How Another's Decisions Shape Our Own. *Journal of Consumer Research*, 50(2), 363–381. <https://doi.org/10.1093/jcr/ucad007>

Perfecto, H., Galak, J., Simmons, J. P., & Nelson, L. D. (2017). Rejecting a bad option feels like choosing a good one. *Journal of Personality and Social Psychology*, 113(5), Article 5. <https://doi.org/10.1037/pspa0000092>

Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika*, 64(2), 191–199.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, 118(1), 42. <https://doi.org/10.1037/a0021150>

Rosenberg, M. (1979). *Conceiving the self*. Basic Books.
<http://books.google.com/books?id=nUJqAAAAMAAJ>

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, 21(4), Article 4.
<https://doi.org/10.3758/BF03197186>

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do Decisions Shape Preference?: Evidence From Blind Choice. *Psychological Science*, 21(9), Article 9.
<https://doi.org/10.1177/0956797610379235>

Shultz, T. R., Léveillé, E., & Lepper, M. R. (1999). Free Choice and Cognitive Dissonance Revisited: Choosing “Lesser Evils” Versus “Greater Goods.” *Personality and Social Psychology Bulletin*, 25(1), 40–48. <https://doi.org/10.1177/0146167299025001004>

- Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. In *The Dynamic Perspective in Personality and Social Psychology* (pp. 283–294). Psychology Press.
- Snyder, M., & Ebbesen, E. B. (1972). Dissonance awareness: A test of dissonance theory versus self-perception theory. *Journal of Experimental Social Psychology*, 8(6), 502–517. [https://doi.org/10.1016/0022-1031\(72\)90076-5](https://doi.org/10.1016/0022-1031(72)90076-5)
- Sokolova, T., & Krishna, A. (2016). Take It or Leave It: How Choosing versus Rejecting Alternatives Affects Information Processing. *Journal of Consumer Research*, 43(4), 614–635. <https://doi.org/10.1093/jcr/ucw049>
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, 89(6), 845.
- Stalder, D. R. (2010). Competing roles for the subfactors of need for closure in moderating dissonance-produced attitude change. *Personality and Individual Differences*, 48(6), 775–778. <https://doi.org/10.1016/j.paid.2010.01.028>
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness* (pp. x, 293). Yale University Press.
- Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2(4), i–109. <https://doi.org/10.1037/h0092987>
- Touré-Tillery, M., & Light, A. E. (2018). No self to spare: How the cognitive structure of the self influences moral behavior. *Organizational Behavior and Human Decision Processes*, 147, 48–64. <https://doi.org/10.1016/j.obhdp.2018.05.002>

- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134(3), 383–403. <https://doi.org/10.1037/0033-2909.134.3.383>
- Vosgerau, J., Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2019). 99% impossible: A valid, or falsifiable, internal meta-analysis. *Journal of Experimental Psychology: General*, 148(9), 1628–1639. <https://doi.org/10.1037/xge0000663>
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), Article 6. <https://doi.org/10.1037/0022-3514.67.6.1049>
- Wedell, D. H. (1997). Another look at reasons for choosing and rejecting. *Memory & Cognition*, 25(6), 873–887. <https://doi.org/10.3758/BF03211332>

All data and materials are available at: <https://osf.io/qhcdy/>