**Action-Inaction Asymmetries in Emotions and Counterfactual Thoughts: Meta-Analysis of the Action Effect   
[Registered Report Stage 1]**

Siu Kit Yeung  
Department of Psychology, Chinese University of Hong Kong, Hong Kong SAR, China  
1155062203@link.cuhk.edu.hk / [yskjdmmh@gmail.com](mailto:yskjdmmh@gmail.com)

^Gilad Feldman

Department of Psychology, University of Hong Kong, Hong Kong SAR, China

[gfeldman@hku.hk](mailto:gfeldman@hku.hk) / [giladfel@gmail.com](mailto:giladfel@gmail.com)

^Corresponding author

Word: abstract – 189, manuscript – 10857

**Author bios:**

Siu Kit Yeung is a Master of Philosophy (MPhil) graduate from the University of Hong Kong psychology department. He is currently a PhD student with the Chinese University of Hong Kong psychology department. His research focuses on Emotions, Judgment, Decision-Making, Mental Health, as well as Open/Meta-Science.

Gilad Feldman is an Assistant Professor with the University of Hong Kong psychology department. His research focuses on Social Psychology, Judgment and Decision-Making, as well as Open/Meta-Science.

**Declaration of Conflict of Interest:**

Several of the studies planned to be included in this meta-analysis were conducted by the authors. Apart from that, we do not have any other conflicts of interest to report.

**Financial disclosure/funding:**

The author(s) received no financial support for the research and/or authorship of this article.

**Authorship declaration:**

Siu Kit Yeung conducted the meta-analysis. Gilad guided Siu Kit Yeung throughout the project and edited the outputs for submission.

**Corresponding author**

Gilad Feldman, Department of Psychology, University of Hong Kong, Hong Kong SAR; gfeldman@hku.hk

**Acknowledgement**

We wrote this meta-analysis with reference to Meta-Analysis Registered Report Templates by Yeung et al. (2021). We appreciate very helpful comments and suggestions from Adrien Fillon and Alison Lam. We are also grateful for important comments and suggestions from the editor Prof. Chris Chambers and peer reviewers Dr. Dan Quintana, Dr. Priyali Rajagopal, and Dr. Emiel Cracco.

## Contributor Roles Taxonomy

|  |  |  |
| --- | --- | --- |
| **Role** | **Siu Kit Yeung** | **Gilad Feldman** |
| Conceptualization | V | V |
| Pre-testing/Simulation | V |  |
| Pre-registration | V |  |
| Data curation | V |  |
| Formal analysis | V |  |
| Investigation | V |  |
| Pre-registration peer review/verification |  | V |
| Literature search |  |  |
| Datafile study/effect coding |  |  |
| Reproducible code (e.g., RMarkdown) | V |  |
| Contacting authors |  |  |
| Data analysis peer review/verification |  |  |
| Methodology | V |  |
| Project administration |  |  |
| Resources |  |  |
| Software | V |  |
| Supervision |  | V |
| Validation |  |  |
| Visualization | V |  |
| Writing-original draft | V |  |
| Writing-review and editing | V |  |

*Note.* Based on Allen and O’Connell (2014). Will be updated in Stage 2. We intend to add another author for Stage 2 for datafile study/effect coding and analyses.

# Abstract

[Note: This is a Stage 1 Registered Report. We will replace the highlighted parts in the Abstract with actual results by Stage 2.]

Action-effect refers to the phenomenon in which people experience, associate, or attribute stronger emotions for action compared to inaction. In this registered report, we conducted a meta-analysis of the action effect literature (*k* = [enter number of studies by Stage 2], *N* = [enter no. of participants by Stage 2], 1982-2021). We found support/no support/mixed support for action-effect in [positive emotions, *g* = X.XX, 95% CI [X.XX, X.XX]], support/no support/mixed support for action-effect in [negative emotions, *g* = X.XX, 95% CI [X.XX, X.XX]], and support/no support/mixed support for action-effect in [counterfactual thought, *g* = X.XX, 95% CI [X.XX, X.XX]]. Study heterogeneity was [low / low to medium / medium / medium to high / high], *Q*(XX) = XXX.XX, *p* = .XXX / < .001, *I²* = XX.XX%. [Summarize results of publication bias tests; to be completed by Stage 2]. Action-effect was stronger [list of conditions in which the effects were stronger, if there was/were, to be entered by Stage 2]. We pre-registered our meta-analysis, with all search protocol, datasets, code, and supplementary made available on the OSF: <https://osf.io/acm24/>.

*Keywords:* Action-effect; Action; Inaction; Meta-Analysis; Registered Report; Regret; Decision Making; Emotions

## 

# Action-Inaction Asymmetries in Emotions and Counterfactual Thoughts: Meta-Analysis of the Action Effect ‎[Registered Report Stage 1]‎

## Introduction

Action and inaction are fundamental aspects of our behavior. One of the most well-known effects in the action-inaction literature is the action-effect, which is the phenomenon that people imagine, associate, or experience stronger emotions for action compared to inaction. It was first demonstrated by Kahneman and Tversky (1982) with a scenario describing two investors who both ended up with negative outcomes following the same investment, with the main difference between the two investors being that one investor switched to that investment from a previous investment (action), whereas the other had considered switching the investment but had finally decided to stick with his original choice (inaction). They found that most people perceived the action investor as experiencing stronger regret compared to the inaction investor, concluding that regret over negative outcomes is stronger when it involved an action decision rather than an inaction decision.

The terms action and inaction have been used in many ways. In the context of the action-effect, action has been used to refer to changing, switching, deviation from past behavioral norm, or doing something, whereas inaction has been used to indicate making no changes, doing nothing, or not doing something (Feldman et al., 2021). Action-effect (Kahneman & Tversky, 1982) has been influential in theoretical developments in the literature covering the emotion of regret (Huang & Zeelenberg, 2012), counterfactual thinking, and norm theory (Kahneman & Miller, 1986). Action-effect has also served as the basis for studying other important phenomena, for example, it has been proposed as a partial explanation for the general preference towards omission over commission as “omission bias”, to avoid possible regret over harm induced (Anderson, 2003; DeScioli et al., 2011). Action-inaction, associated emotions, and cognitions have been shown to be important in understanding human judgment and decision making in various domains, such as investment (Kahneman & Tversky, 1982), health-medicine (Brewer et al., 2016), career or education (Gilovich & Medvec, 1994), sports (Bar-Eli et al., 2007), and morality (Jamison et al., 2020).

Over the years, researchers studying the phenomenon have identified moderators and boundary conditions that have been shown to weaken or even reverse the action-effect, raising the need for a systematic review meta-analysis of the effect and the constraints over its generalizability (e.g., Gilovich & Medvec, 1994; Zeelenberg et al., 2002). Scholars have also suggested several ways to improve various aspects of empirical studies demonstrating the phenomenon and pushing the literature forward. For example, the literature would greatly benefit from better clarity regarding the definitions of the terms action and inaction and linkage with the concepts of normality and exceptionality (Feldman & Albarracín, 2017; Feldman et al., 2020, 2021), a comparison of effects using different methodology and context (Zeelenberg et al., 2002), measurements, and study designs (N’gbala & Branscombe, 1997; Zhang et al., 2005).

The present meta-analysis registered report aims to contribute to the literature by conducting a systematic quantitative review of the literature on the action-effect to further our theoretical and empirical understanding of the phenomenon with categorizations of effects in the literature, assessment of aggregated effects, and the identification and quantification of moderators.

We begin with an introduction of the classic pioneering study by Kahneman and Tversky (1982) and a brief review of the theoretical explanations for the action-effect, followed by a review of the key findings and debates in the literature, identifying potential moderators.

## Action-effect

Action-effect was first demonstrated by Kahneman and Tversky in 1982 and has been immensely influential ever since. At the time of writing (August 2022), we identified 2626 citations of the article (according to Google Scholar), with many important follow-up theoretical and empirical articles, with the most well-known being Kahneman and Miller’s (1986) norm theory.

Kahneman and Tversky (1982) presented 138 participants with scenarios describing two investors: George, who decided to switch his investment, and Paul, who decided not to change his initial investment with both being described as having experienced the same negative outcome. Participants were then asked who they thought would experience stronger regret. They found that most participants perceived that action-George would experience stronger regret than inaction-Paul. Many follow-up studies have successfully replicated Kahneman and Tversky’s (1982) findings (e.g. Connolly et al., 1997; Fillon et al., 2022; Gleicher et al., 1990, Landman, 1987; Sepehrinia et al., 2022). For example, one of the very few action-effect studies measuring both pleasant and unpleasant emotions by Landman (1987) found support for action-effect for both elation and regret across different scenarios.

In this meta-analysis, we focused on action-inaction asymmetries regarding emotions and counterfactual thoughts. This is different from other related action-inaction judgments, decision making, and social psychology effects, such as studies of moral judgments or decisions (for a meta-analysis on omission bias, see Yeung et al., 2021), status quo bias, and default effect (Feldman et al., 2020).

**What are action and inaction?**

Action effect has laid the foundation for discussing action and inaction, yet the terms action and inaction were never clearly defined until recently (Feldman et al., 2021). When referring to “action” or describing an “action”, some (e.g., Kahneman & Tversky, 1982) may refer to change or deviation from past behavioral norms, yet this does not always have to be the case. Sometimes action (e.g., Gilovich & Medvec, 1994) is used to refer to “doing something”, for example, going to college, whereas inaction refers to “not doing something” or “doing nothing”, for example, not going to college. “Action” may also mean physically active or making a proactive choice, whereas “inaction” may mean physically inactive, making a passive choice (e.g., avoidance), or procrastination (Feldman et al., 2021). However, most studies on the action-effect seemed to either use the “change vs no change” meaning or the “doing something vs not doing something” meaning. Apart from the recent work by Sepehrinia et al. (2022) which manipulated these two types of action-inaction operationalizations in a single study, we are unaware of any study comparing action-effects with two operationalizations *within* a study (but see Yeung & Feldman, 2022 for related discussions regarding discrepancies of findings *between* studies). The unclear meanings have been identified as a major gap in the literature leading to confounded usage by Feldman et al. (2021), and so in our meta-analysis, we aim to try and carefully disentangle the different types and the possible similarities and/or differences in effects.

## Norm Theory and Normality

Action-effect is related to the concept of normality and can be explained by norm theory (Kahneman & Miller, 1986). Normality refers to the extent to which behaviors, circumstances, or outcomes are perceived to be normal. Normality may depend on cognitive availability of events or stimuli, and it is likely related to counterfactual thought – retrievability of related stimuli or events and imagining alternative realities (Feldman et al., 2020). The term exceptionality is often used to refer to the deviation from a reference point, such as past behavioral norms (Feldman et al., 2020; Fillon et al., 2020). Regret appears to be stronger when the behavior before the outcome is more exceptional, and the argument was that is it easier to imagine normal alternatives (routine behavior) that may have resulted in a more favorable outcome (Feldman, 2020). Regret and upwards counterfactuals are positively correlated, though the causal direction remains unclear (Coricelli & Rustichini, 2010; Feldman, 2020; Fillon et al., 2020). Norm theory suggested that in action-effect situations, inaction seems more normal than action, and action is associated with more counterfactual thoughts. It therefore seems easier to imagine abstaining from action following the established norm than to imagine action that deviates from that norm. Thus, people imagine, attribute, or experience stronger regret following action compared to following inaction. For example, in the context of the Kahneman and Tversky (1982) investor scenario, action means switching or deviating from a previous investment decision (past behavioral norm as the reference point), the reference point. In the context of risky financial investments that may result in negative outcomes, actions to make an investment or to change investments may be perceived as less normal, or as a deviation from one’s past behavioral norm. That being said, we note that “action” does not necessarily mean a change or a deviation from one’s past behavioral norm, especially in studies where the experimenters asked participants to think of “something they did” (e.g., Study 1 and Study 5 of Gilovich & Medvec, 1994). The norm-theory account may not be applicable to those situations.

As mentioned above, there is likely a positive correlation between regret and counterfactual thoughts. It had been assumed that higher counterfactual mutability of actions compared to inactions accounts for the action-effect (Kahneman & Miller, 1986), but very few studies measured both counterfactual thoughts and regret simultaneously (Fillon et al., 2020; N’gbala & Branscombe, 1997). In follow-up studies testing the norm theory account of action-effect, N’gbala and Branscombe (1997) and Davis et al. (1995) failed to find support for the proposed mediating role of counterfactual thought. Davis et al. (1995) compared counterfactual mutability of action vs inaction for spouses and parents who experienced accidental traumatic events and found that people mutated inactions more than actions. N’gbala and Branscombe (1997) argued that “There are situations where actions and inactions can be perceived as equally mutable: that is when both can be construed as instrumental or causal for achieving the desired outcome.” (p. 328). That said, it is premature to conclude that there are no differences in counterfactual thought between action and inaction based on one or two articles, and so a meta-analysis on the action-inaction asymmetries of counterfactual thoughts could help assess this systematically. We also note that there has been evidence, with larger sample sizes and across several studies, suggesting action is associated with more “what if” thoughts compared to inaction (Byrne & McEleney, 2000). The discrepancies in findings between studies may be due to differences in measures, scenarios, or study design. One possible explanation suggested by Byrne and McEleney (2000) is that there may be a stronger effect with within-subject design.

## Decision Justification Theory

Another key theory in the domain of action-inaction is the Decision Justification Theory (DJT; Connolly & Zeelenberg, 2002), which focuses on the justifiability of decisions resulting in regret. Linking the two theories, exceptional decisions may be more likely to be perceived as less justified than normal decisions, thereby resulting in stronger regret when things go wrong (Reb & Connolly, 2010). This is because it is easier to justify routine or normal decisions to oneself and to others compared to exceptional or abnormal decisions (Reb & Connolly, 2010). In the context of the two-investor scenario, action (deviating from past personal behavioral norms) may be perceived as more exceptional and less justified than inaction (sticking with the same option). However, under some situations and some social environments, action (in the form of change) may be perceived as more normal and more justified than inaction (in the form of no change), notably but not limited to situations with negative prior outcomes and problems or certain situational expectations (Bar-Eli et al., 2007; Connolly & Zeelenberg, 2002; Feldman, 2020; Feldman & Albarracín, 2017; Olsen, 2017; Zeelenberg et al., 2002). Action norms in some situations may partly explain findings of inaction-effect, of stronger regret for inaction than for action (e.g., Feldman & Albarracín, 2017; Zeelenberg et al., 2002). Given prior negative outcomes or situational expectations for taking action, it would be more justifiable to act in order to tackle the problem, thereby leading to stronger regret if one failed to act.

## Additional theoretical accounts

There are other theoretical explanations of the action-effect. For example, action is associated with stronger responsibility than inaction and is perceived as more agentic (Connolly et al., 1997; Yeung et al., 2021). Action would be associated with stronger regret than inaction, given that someone would perceive or experience a stronger sense of responsibility for the negative outcome resulting from an action (Connolly et al., 1997). Another explanation is that action is generally perceived as more effortful, and the actor is perceived to experience stronger regret over wasted effort (N’gbala & Branscombe, 1997). N’gbala and Branscombe (1997) found that the actor is perceived as less wise as the actor wastes too much effort and time. It seems plausible that wisdom is likely positively associated with justifiability (Garrett, 1996), but we are not aware of studies testing both in action-effect studies. We note these accounts, but these explanations are outside the scope of this meta-analysis.

We summarized the theoretical accounts of action-effect in Table 1. These theoretical explanations remain under debate and are facing multiple challenges. For example, some elements of norm theory, such as counterfactual mutation (Kahneman & Miller, 1986), have received no support from some studies (Davis et al., 1995; N’gbala & Branscombe, 1997). In these theoretical accounts, constructs such as action, inaction, normality, exceptionality, and justifiability were not clearly defined and explained, and the expected relationships between these constructs given theory were not clearly elaborated (Feldman, 2020; Feldman & Albarracín, 2017; Feldman et al., 2021). Theories can improve in outlining testable hypotheses that would favor one account over the other. However, these theories are not necessarily contradictory and may jointly explain the phenomenon. An action may be regretted more because of several different reasons - it is perceived as abnormal, it is perceived as more blameworthy or responsible, it is perceived as wasted effort, or it is perceived as unjustified. It is possible all these factors contribute to action-effect in some situations, but one of these accounts may be more relevant in some situations. That said, very limited studies have compared these different contributing factors of action-effect and there is much need for more research on the intersection of the theoretical paradigms. Nonetheless, even though still lacking in sound theoretical and construct specificity, action-effect as a phenomenon seems to have empirically strong foundations.

Table 1

*Theoretical explanations of the action-effect*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Norm Theory | Decision Justification Theory | Action-  responsibility link | Action-effort link |
| Focus and assumptions | Focuses on normality and exceptionality of action and inaction. Exceptional behavior is more strongly regretted than routine behavior. Assumes inaction to be more normal than action. | Focuses on justifiability. Less justified decisions are regretted more than more justified decisions. | Action is attributed more blame than inaction. Stronger regret due to stronger assumed responsibility. | Action is perceived as more effortful, and as wasted effort if the outcome is negative. |
| Implications regarding action-effect | When inaction is more normal (e.g., past behavior, social norms, etc.), inaction may be regretted more than inaction. | When inaction is more justified (e.g., prior positive outcomes, situational expectations for inaction), inaction may be regretted more than inaction. | Action is regretted more than inaction | Action is regretted more than inaction in a within-subject design. |
| Evidence and empirical challenges | 1) Assumptions regarding whether inactions or actions are more normal are typically untested  2) N’gbala and Branscombe (1997) failed to find support for stronger counterfactual mutation of action  3) Byrne and McEleney (2000) found support for stronger “if only” thoughts for action compared to inaction. | 1) Assumptions regarding whether inactions or actions are more justifiable are typically untested.  2) Unclear link between normality and justifiability. | 1) Supported by a recent meta-analysis (Yeung et al., 2021).  2) Boundary condition: under certain social roles in which the decision-maker is responsible for the target, such effect is minimized (Haidt & Baron, 1996). | 1) No direct evidence regarding effort. Action is perceived to be a waste of energy and time (N’gbala & Branscombe, 1997)  2) Action-effect seems stronger in within-subject designs  (N’gbala & Branscombe, 1997; Zhang et al., 2005). |

## Boundary conditions of Action Effect

Sometimes action-effect can be weakened or even reversed into inaction-effect. Identified moderators in the literature so far include study design (N’gbala & Branscombe, 1997; Zhang et al., 2005), prior outcomes (Zeelenberg et al., 2002), temporal distance (Feldman et al., 1999; Gilovich & Medvec, 1994), generality (Davison & Feeney, 2008), and study realism (Gilovich & Medvec, 1994). For example, Gilovich and Medvec (1994) conducted five studies on the temporal pattern of the action effect using different methods including telephone surveys, two-scenario experiments, and questionnaires. They found that participants tended to both evaluate others’ regrets and report their own regrets as being stronger for inaction compared to action, the more distant in the past the experience was. Using the Kahneman and Tversky (1982) investor vignette, they showed a reversal of the action-effect by manipulating temporal distance.

Recently, Yeung and Feldman (2022) conducted replications of Gilovich and Medvec (1994) and found support for the temporal reversal of action-effect in the long run with scenario experiments, yet failed to find support for an action-effect or an inaction-effect in recalled short-term and long-term real-life experiences. They summarized it was unclear why they observed discrepancies between these studies since it could be more than just the contrast between scenarios and recalls. Another possibility raised was that these may be due to the different meanings of action and inaction used in these studies. Namely, in the vignette replications action meant change and inaction meant no change, whereas in the recall replications action meant doing something and inaction meant not doing something. Other differences between the studies were about actual first-person experiences versus evaluations of third-person emotions (Yeung & Feldman, 2022), and methodological/scenario constraints on generality (Byrne & McEleney, 2000; Simons et al., 2017). Regardless, these findings suggest that action-inaction-associated regret is a much more complex phenomenon than initially assumed.

Zeelenberg et al. (2002) showed a reversal of action-effect using prior outcomes. They showed that with prior positive and neutral outcomes, the action-effect findings from Kahneman and Tversky (1982) held. However, they showed an inaction effect when prior outcomes were negative, suggesting that these shift expectations for agents to act attempting to resolve the problem. A similar demonstration by Feldman and Albarracín (2017) showed similar findings on action-inaction social norms as moderators of action-effect. Past behavior, expectations, and social norms serve as reference points that people use to assess what is considered normal and justifiable (Feldman, 2020). A meta-analysis allows for systematic quantification of the effect, investigating possible moderators.

## Aim and scope

In this meta-analysis, we investigated action-inaction asymmetries of emotions and counterfactual thoughts. We included counterfactual thoughts as those have been shown as being associated with emotions. We focused on these dependent variables as these were the first initial demonstrations of action-inaction asymmetries, and based on our experience in conducting replications in this domain these dependent variables seem to be the most commonly studied in the action-inaction domain.

The current project will not cover action-inaction asymmetries regarding moral judgments and decisions. Some of those asymmetries have already been investigated in other meta-analyses, such as in the meta-analysis on omission bias by Yeung et al. (2021). Omission bias and action-effect are related but distinct, with different dependent variables, moderators of interest, and theoretical paradigms (see Feldman et al., 2020; Yeung et al., 2021 for more details). Norm theory is a framework that has been used to try and align omission bias with action-effect (Feldman et al., 2020) yet this remains to be studied further, and there are other theoretical explanations for both effects (decision justification theory, action-effort link, action-causality-link, etc.) that seem more suitable for explaining one effect yet not the other.

We aimed to address the following research questions: 1) What is the overall empirical support for the action effect? 2) What factors moderate the action-effect? 3) Under what conditions are action-effect weakened or reversed into inaction-effect?

### Action-effect: Main effect

Our first aim was to examine the overall effect of action-effect. We expected the evidence to be in support of action-effect with a positive effect meaningfully different from null (null not included in confidence intervals) (H1overall). Specifically (see Table 2 for term elaboration):

H1a: Action is associated with stronger negative emotion than inaction, given negative outcome.

H1b: Action is associated with stronger positive emotion than inaction, given positive outcome.

H1c: Action is associated with more counterfactual thought (number of counterfactuals) than inaction.

A recent study by Fillon et al. (2022), one of the very few studies that investigated both regret and joy, found stronger effects for regret and weaker effects for joy, possibly because of negativity bias, meaning that bad seem to be perceived to be more agentic (Feldman et al., 2016) and has stronger emotional impact than good (Baumeister et al., 2001; Kahneman & Tversky, 1979). Their findings were different than the more classic Landman (1987) and Sepehrinia et al. (2022) who failed to find support for differences between joy and regret. Given such mixed findings, we set this direction as exploratory. The vast majority of the studies we know from the literature only investigated action-effect for negative emotions, with very few studies measuring positive emotions (e.g., Fillon et al., 2022; Landman, 1987). If we find 5 studies for action-inaction asymmetries in positive emotions, then we will conduct exploratory moderator analyses comparing positive emotions studies and negative emotions studies.

Exploratory hypotheses (if 5 studies or more on positive emotions): Action-effect is stronger/weaker for negative emotions compared to positive emotions.

Another exploratory direction is regarding possible differences between upward counterfactuals and downward counterfactuals. To the best of our knowledge, the vast majority of studies measured upward counterfactuals, but there might be studies on downward counterfactuals.

Therefore, if we find at least 5 studies on action-inaction asymmetries on downward counterfactual thoughts (and at least 5 studies for upward counterfactual thoughts), then we would conduct exploratory analyses for moderation of upward counterfactuals versus downward counterfactuals.

Exploratory hypotheses (if 5 studies or more on downwards counterfactuals): Action-effect is stronger/weaker for upwards counterfactuals compared to downwards counterfactuals

Table 2

*Clarifications regarding terms used in the context of the action-effect*

|  |  |
| --- | --- |
| Terms | Description with examples |
| Associated with | 1. Type 1: Perception (e.g., Kahneman & Tversky, 1982) in which more participants perceived that the investor (a third person) who switched would experience stronger regret. 2. Type 2: Actual experience (e.g., Gilovich & Medvec, 1994 Study 1), in which more participants experienced stronger regret for inaction than action when retrospectively thinking about their lifetime regrets. 3. “Action is *associated* with stronger emotions than inaction” means that most people perceive and/or experience stronger emotions acting compared to not acting. |
| Counterfactual thoughts | Thoughts that are “counter to the factual events”, or possible alternatives of what if the decision-maker chooses another option.  For example, in N’gbala and Branscombe (1997), the experimenter asked participants to imagine and describe their “what if” thoughts. After that, researchers coded the responses as mutations to action or mutations to inaction. |
| Positive outcome  Negative outcome | The consequence for the decision-maker is desirable. For example, receiving a better grade in Landman (1987).  The consequence for the decision-maker is undesirable. For example, losing money in Kahneman and Tversky (1982) |
| Positive emotion | Pleasant emotions including but not limited to satisfaction, happiness, joy, elation, feeling good, feeling better, relief. For example, Landman (1987) asked participants to compare who (the actor vs the non-actor) would feel better. |
| Negative emotion | Unpleasant emotions including but not limited to guilt, regret, disappointment, sadness, feeling bad. For example, Kahneman and Tversky (1982) asked participants to compare regret for the actor and the non-actor. |

### Confirmatory Moderators

#### Temporal Distance

Does temporal distance moderate the action-effect? Gilovich and Medvec (1994) replicated action-effect in the short-term yet found support for temporal distance as a moderator. Weaker or reversed action-effect in the long-term may be due to cognitive accessibility (Rajagopal et al., 2006), or perhaps because the consequences of actions occur more quickly and tend to be more limited, but the consequences of inactions may occur later (Gilovich & Medvec, 1995). Another possible explanation has to do with the type of affect elicited by temporal distance (Gilovich et al., 1998).

Evidence regarding this moderator is mixed. Bonnefon and Zhang (2008), Byrne and McEleney (2000), Feldman et al. (1999), and Towers et al. (2016) failed to find support for the temporal pattern of action-inaction regret. Recently, Yeung and Feldman (2022) replicated Gilovich and Medvec (1994) and found support for action-effect in the short-term and inaction-effect in the long-term with college decision scenario experiments yet failed to find support for the temporal effect with real-life experience studies. Null findings do not necessarily mean that the moderation effect of temporal distance does not exist, as these may be caused by various methodological factors (Feldman et al., 1999), yet it raises the need to systematically examine this evidence. To the best of our knowledge, the vast majority of studies for temporal pattern of action-effect measured negative emotions but not positive emotions, and we therefore do not have any a-priori expectation regarding such temporal pattern for positive emotions. We use the number of years[[1]](#footnote-2) as the unit to code the temporal distance. We also conduct such analyses categorically (recent or current events – less than a year, more distant events – a year ago or more, lifetime events). Therefore, we hypothesize the following:

H2: Temporal distance moderates the relationship between action and inaction on emotions. The more distant[[2]](#footnote-3) the behavior is in the past, the weaker the action-effect. Action-effect is stronger for recent or current events, compared to more distant (1 year ago or more) or major lifetime events.

#### Study Design: Between versus within versus comparison

Researchers initially studied action-effect using within-subject or one-sample comparison designs (N’gbala & Branscombe, 1997; Zhang et al., 2005), yet there is evidence that action-effect is weaker using between-subject designs (N’gbala & Branscombe, 1997).

Comparison studies (where the dependent variable is a choice between action or inaction) and within-subject studies (where the dependent variables are two separate ratings of action and inaction) allow for an evaluation of action against inaction and therefore may provide better control for contextual factors and confounding variables. Yet, some have argued that effects in within and comparison designs may be the result of stronger demand effects. The argument is that participants may be able to guess the hypothesis of the study and answer in a way they think would best fit with the experimenter’s expectations (Charness et al., 2012; Kahneman & Tversky, 1982b), though recent studies found little support for that argument (Lambdin & Shaffer, 2009; Mummolo & Peterson, 2019).

Another argument in this debate is that between subject designs are more realistic. In real-life, people more often face situations in which they are only presented with either action or inaction resulting in a negative outcome, rather than both together. However, between-subject designs often lack a meaningful reference point to be used for comparison (Zhang et al., 2005). This may result in arbitrary choices of some other contextual cues to compare against.

We expected to find support for an action-effect, regardless of design, yet weaker effects in between designs compared to comparison studies and within-subject studies. We also expected a stronger effect using comparison designs compared to within-subject designs and between-subject designs. Fillon et al. (2020) argued that comparison designs create an artificial contrast leading to an overestimated effect size for exceptionality bias, a related yet distinct judgment and decision-making bias. They found effect sizes in comparison studies were over three times larger than those in experimental studies.

Therefore:

H3a: H1 will be supported across between, within, and one-sample comparison designs (confidence intervals do not overlap with the null).

H3b: Action-effect is stronger using within-subject designs compared to between-subject designs.

H3c: Action-effect is stronger using comparison designs compared to within-subject designs.

[H3b+H3c: Action-effect is stronger using comparison designs compared to between-subject designs.]

#### Normality: Prior outcomes and social norms

Zeelenberg et al. (2002) found that given negative prior outcomes, participants associated more regret with inaction compared to action, possibly due to an expectation or formed norm for an action to change so that future losses would be avoided (Feldman & Albarracín, 2017). This is in line with decision justification theory (Connolly & Zeelenberg, 2002), arguing that decisions (action given negative prior outcomes) that are better socially justified are regretted less. Most studies demonstrating the action-effect (e.g., Kahneman & Tversky, 1982) adopted isolated decision scenarios, in which there was no background information regarding the outcomes of prior decisions.

We are interested in investigating if there would be a substantial difference in effect size and a possible reversal of the sign with prior negative outcomes compared to having no prior outcomes, neutral prior outcomes, or positive prior outcomes.

H4a*:* H1 will be supported given no prior outcomes, or with neutral or positive prior outcomes.

H4b: Action-effect is moderated given negative prior outcomes (weaker or reversed into an inaction effect).

Based on norm theory (Kahneman & Tversky, 1982), people experience stronger regret for abnormal decisions and weaker regret for normal decisions. Feldman and Albarracín (2017) and Feldman (2020) proposed that there are three types of normality: past behavioral norm, social norm, and situational/role norm, and that given no information about normality and under inaction norm, action-effect is more likely to occur. In contrast, under action norm, action-effect may be weakened or even reversed into inaction-effect. Therefore:

H4c: Action-effect is stronger under inaction norms compared to action norms.

H4d: Action-effect is stronger given no norm information compared to action norms.

H4e: Action-effect is stronger under inaction past behavior compared to action past behavior.

H4f: Action-effect is stronger given no mention of past behavior compared to action past behavior.

#### Specific vs General

Gilovich and Medvec (1995) proposed that action regrets are more likely to be based on decisions at a specific time, and inaction regrets are more likely to be about accumulated or generalized events. To test this idea, Davison and Feeney (2008) examined regret from the perspective of autobiographical memory, and found that people experienced more inaction regret for general events, with elderly participants and participants at 40s, but failed to find support for a meaningful difference between general memories and specific memories for action-regret. Their findings were consistent with studies that demonstrated inaction effects for general life events (Gilovich & Medvec, 1994; Yeung & Feldman, 2022). However, there appear to be very limited follow-up studies or replications with other samples. Therefore:

H5: Action-effect is stronger for specific events compared to general events.

#### Hypothetical vs real-life experience

Some studies presented participants with hypothetical vignette scenarios (e.g., Kahneman & Tversky, 1982; Landman, 1987) whereas other studies (e.g., Feldman et al., 1999; Gilovich & Medvec, 1994) tested autobiographical real-life experiences of action-inaction emotions. Real-life events are different from well-controlled hypothetical scenarios with the same outcome, as negative outcomes from action and inaction may not be equal in strength, resulting in systematic variations between action and inaction experiences (Feldman et al., 1999). There appear to be more successful demonstrations of action-effect with hypothetical scenario studies (Gleicher et al., 1990; Kahneman & Tversky, 1982; Landman, 1987), but more mixed findings, null findings, or even reversed findings with real-life experience studies (Feldman et al., 1999; Gilovich & Medvec, 1994; Yeung & Feldman, 2022). The mechanism is unclear and may be because inaction regrets in real-life (e.g., not going to college) tend to be longer-lasting and result in more consequential outcomes (Gilovich & Medvec, 1995). The differences between hypothetical scenarios and real-life events may be confounded by the fact that most hypothetical scenarios seem to focus on short-term scenarios, whereas most real-life events studies seem to focus on lifetime or long-term experiences. Temporal distance may be a potential moderator. Another explanation proposed by Feldman et al. (1999) is that “if it is easier to anticipate the potential for harm due to action than due to inaction, people might be more

conservative in their choice of which actions to take. Consequently, greater harm and more intense regrets would result from failures to act” (p. 254). Therefore:

H6: Action-effect is stronger for studies using hypothetical scenarios than for real-life events.

### Exploratory Directions

#### Target: Self versus others

Some studies asked participants to imagine, attribute and/or compare the level of emotion of others (e.g., Kahneman & Tversky, 1982 Action George vs Inaction Paul), whereas some studies asked participants to rate or compare regret for action vs inaction for self (e.g., Gilovich & Medvec, 1994). Self-other distinction is often a key moderator in judgment and decision-making research. A recent meta-analysis on omission bias (Yeung et al., 2021) investigated self-other differences and found mixed results with different models yet seemed to generally suggest omission bias is stronger if the target is self, compared to when the target is others. Based on the findings from the omission/action bias literature (Zikmund-Fisher et al., 2006), it seems plausible to expect a stronger action-effect for self compared to for others.

#### Meanings of action and inaction

In the classic Kahneman and Tversky (1982) demonstration of action-effect, “action” meant changing from a prior decision (deviating from past behavioral norm) whereas “inaction” meant persisting with a prior decision. However, over the years there have been many other uses of the terms action and inaction. See our discussion in the introduction. We planned to compare action-effect given meanings of “change vs no change” (example: Kahneman & Tversky, 1982, which seem to be more related to norm theory and the concept of normality, see Yeung & Feldman, 2022), “doing something vs not doing something/failure to do something” (example: Gilovich & Medvec, 1994 Study 5), and other meanings (depending on what meanings we find in the literature). If there are too few studies (under five) using any of the meanings, we would not analyze those meanings. It seems plausible that action-effect would be stronger if action involves changing from one investment target/school/company to another one, compared to when action does not involve switching (Sepehrinia et al., 2022). However, since we are only aware of one study (Sepehrinia et al., 2022) directly comparing these two operationalizations and they found support for differences for two domains but failed to find evidence for two other domains, we do not have specific hypotheses and aim to examine these as an exploratory moderator of action-effect.

# Methods

[Note: Written in past tense to demonstrate methods section after completion, but has yet to be conducted. Will be filled and updated after pre-registration and data collection.]

## Open Science Disclosures

We shared all procedures, materials, datasets, and code on the Open Science Framework: <https://osf.io/acm24/>.

We conducted an initial unstructured pre-search to construct, test, and refine our search syntax. We then posted notices on listservs asking authors to alert us of possible related articles and unpublished studies: Society for Judgment and Decision Making (SJDM), European Association for Decision Making (EADM), European Association of Social Psychology (EASP), Society for Personality and Social Psychology (SPSP), and Social Psychology Network (SPN) on [Dates to be inserted in Stage 2]. [Note: The list will be updated in Stage 2 if changed]

We provided a notice template in the Template for Contacting Authors for Published and Unpublished Data on Listservs section of the Supplementary. Systematic data collection has not been conducted for this project. There are no other unreported pre-registrations for this meta-analysis project. See Open Science Disclosures in Supplementary for details.

## Literature search

To find articles relevant to our topic, we used Google Scholar (Walters, 2007), PsychINFO, Scopus, ProQuest Dissertations and Theses Global, Web of Science (Moreau & Gamble, 2020), and Opengrey.eu. We used Boolean Logic operators such as “OR”, and “AND” in the search pattern to connect action-inaction/commission-omission and positive and negative emotions. Table 3 lists all search patterns. We validated and pre-tested the search pattern with 10 randomly selected articles out of all articles related to action-effect mentioned in this manuscript.

At Stage 2, we reran searches at least two times to ensure we are up to date with the literature. All database searches achieved XX hits. The date last searched was [to be inserted in Stage 2]. We used Zotero, a referencing software, to deduplicate. After removing duplicates, a total of XX published articles, unpublished articles, and datasets were initially identified and downloaded from the primary database search.

After that, a search for relevant papers not listed in the primary database search was conducted, by searching for papers listed under the “related articles” and “cited by” features in Google Scholar using the identified list of articles. The date last searched was [to be inserted in Stage 2]. We also conducted one additional round of search by scanning the reference sections of identified articles from our primary search. The date last searched was [to be inserted in Stage 2]. The outcome was a total of YY articles.

Furthermore, we identified authors in the field of the action-effect literature, searched through their available papers, and contacted them along with authors of other identified articles to ensure full coverage, and to maximize inclusion of unpublished data and/or manuscripts that are also relevant (Feltz & May, 2017). See Supplementary Materials template for contacting authors subsection. We contacted authors of studies with missing data for clarifications and information. We documented this process and the relevant results in the coding sheet (Contacting the Authors tab). We saved all final studies included in the total search into a cloud folder, accessible from the OSF link or directly via: [to be provided by Stage 2].

In total, we contacted [number of authors, to be entered by Stage 2], [number of authors responded, to be entered by Stage 2] responded, and [number of authors that provided additional relevant papers, data, or information, to be entered by Stage 2] provided [number of extra studies included through this search process, to be entered by Stage 2] additional data/papers that were eventually included in our meta-analysis (Appelbaum et al., 2018). Lastly, we issued a call for unpublished findings on online forums, research platforms, and social media (e.g., ResearchGate, listservs, social media) on [insert dates]. We set up a project on ResearchGate and added all identified articles as references, where possible, as another method of trying to reach and notify authors about this project. We used ResearchGate and OSF to keep track related of articles, projects, and authors, and to provide an open-access list of available studies (links: [insert link]). The OSF project was also used to store the Datafile (with all articles and studies), the RMarkdown code and outputs, and preprints.

After the above search procedures, the lead author and a coauthor scanned all abstracts, tables, and method sections to identify the relevance of the sources. If the article was relevant to our analysis, the lead author and a coauthor read more of the article to determine whether it met the inclusion criteria. This round of screening enabled us to exclude XX articles, reducing our sample of studies to YY articles with a total of YYYY participants. [This is planned and depends on identifying another coder. Will be updated if changed.]

Table 3

*Search syntax, date of searches, and number of results returned for each database used in meta-analysis*

|  |  |  |  |
| --- | --- | --- | --- |
| **Database** | **Date of Coverage** | **# Results** | **Search syntax** |
| PsychINFO | [To be inserted by Stage 2] | [To be inserted by Stage 2] | ("action effect\*" OR "actor effect\*" OR "inaction effect\*" OR ‎‎"Kahneman Tversky 1982" OR "action\*" OR "inaction\*" OR “omission\*" OR "commission\*” OR "chang\*" "switch\*") AND ...  Combination 1 with negative emotions:  AND (“regret\*” OR "sad\*" OR "disappoint\*" OR “remorse\*” OR “guilt\*” ‎ OR "anger\*" OR “frustration\*” OR ‎‎“irritation\*” OR “shame\*” OR “bad\*” OR "negativ\*")  Combination 2 with positive emotions:  AND (“happ\*” OR “joy\*” OR “good\*” ‎OR “elation\*” OR “pride\*” OR “satisf\*” OR "positiv\*")  Combination 3 with ‎counterfactual thought: ‎ AND ("counterfactual\*” OR “if I had\*” OR “what if\*” OR "alternative realit\*")‎  Combination 4 with generalized emotions words to be checked and categorized after reading:  AND (“feel\*” OR “emotion\*”) |
| Scopus | [To be inserted by Stage 2] | [To be inserted by Stage 2] | Same or variation of the above. |
| ProQuest Dissertations & Theses Global | [To be inserted by Stage 2] | [To be inserted by Stage 2] | Same or variation of the above. |
| Web of Science | [To be inserted by Stage 2] | [To be inserted by Stage 2] | Same or variation of the above. |
| Opengrey.eu | [To be inserted by Stage 2] | [To be inserted by Stage 2] | Same or variation of the above. |
| Google Scholar | [To be inserted by Stage 2] | [To be inserted by Stage 2] | Same or variation of the above. |

## Inclusion and exclusion criteria

We detail several inclusion and exclusion criteria.

First, the studies had to include adequate statistical information for computing the effect size for the action-inaction difference in intensity of emotions. In cases of missing statistical data, we first attempted to contact the authors. When plots are provided, we would attempt to use plot digitizers tools such as Web Plot Digitizer (Drevon et al., 2017), GraphClick (Arizona Software Inc., 2010), DataThief III (Tummers, 2006; also see Flower et al., 2016 for reliability and validity of GraphClick and DataThief III), or digitize R package (Poisot et al., 2016) to extract data from plots. We will document the methods used for obtaining required statistics in the coding sheet. When we were not able to obtain the required statistics or raw data from the article or through emailing the authors, we excluded the studies, even if the studies met all other search criteria. We excluded all non-experimental studies.

Second, we excluded articles not written in English, unless we successfully obtained data or information needed for coding designs and effects. This means that in the initial screening, if the title is in English or if we understand that the title is related to action-effect, we will attempt to contact the authors to provide data.

Third, we included both published and unpublished studies, from 1982 to [insert date when the search is officially completed, in Stage 2]. We chose 1982 to be the cut-off year as the first study on the action-effect by Kahneman and Tversky was published in 1982.

Fourthly, we only included studies with a clear explicit contrast between action and inaction.

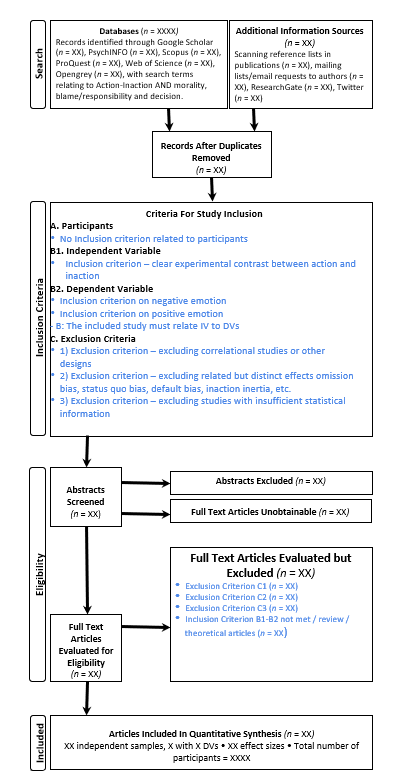
## Screening

Studies collected through database searches and through contacting authors were assessed for their eligibility based on their titles, abstracts, and contents. A collaborator ([insert name of coauthor by Stage 2]) joined the team after receiving an in-principle acceptance to help with search, screening, and coding together with the lead author [will be updated in Stage 2 or if recruitment was not successful].

We first discussed and developed a screening procedure and script (adapted from Polanin et al., 2019). We looked for keywords, notably action/commission and inaction/omission, emotions, and words indicating experimental designs. We omitted articles about other similar but distinct action-inaction-related effects, correlational studies, cross-sectional studies, qualitative studies, review papers, and papers published before 1982. Through these discussions, we ensured all authors agree on and understand the definitions and meanings of all items in the screening tool. There were some borderline cases of inclusion vs exclusion. After our discussions on these borderline cases, we ensured that we maintained and implemented our criteria of inclusion vs exclusion consistently. When the two coders (Siu Kit Yeung and [insert name of coauthor by Stage 2]) could not reach an agreement on certain inclusion/exclusion, the corresponding author (Gilad Feldman) would make the final decision as a moderator. We documented and explained all decisions for inclusion and exclusion clearly, transparently, and systematically within the Action-effect-meta-excel-v1.xlsx file Article list Inc+Exc Criterion tab. We provided the details of articles/studies excluded at the screening stage and eligibility stage in the Supplementary Open Science Disclosures section and the Search Flow Diagram below.

Figure 1

*Search Flow Diagram*



*Note.* We adapted the above diagram based on Moreau and Gamble (2020).

We saved all preliminary studies included in the total search into a cloud folder. The full-texts are accessible from the project OSF link or directly via [to be added at Stage 2]. We scanned all articles to determine whether we should include them in the main coding sheet or not.

## Coding and pre-testing

We developed a data coding file to keep a record of our decisions at different stages (Arslan, 2019; Obels et al., 2020; Siddaway et al., 2019). Before we began with the coding process, we pilot-tested 10 randomly-selected studies in two stages and made needed adjustments in every stage. The lead author completed the coding process for the pretests, and a coauthor verified. We documented all disagreements and reported decisions. Coders discussed and documented disagreements and adjusted the coding accordingly. This pretest coding clarified the definitions and categorization of the moderator categories, to reduce potential disagreements during the actual coding stage.

## Coding

Once we completed the pre-test coding and confirmed the included studies, the lead author coded the studies independently, then a coauthor verified the coding. The two discussed disagreements. When the two failed to reach an agreement, a moderator made the final decision.

**Variables and Design in the studies**

We included studies with continuous variables and categorical variables measures. Studies with continuous variables adopt Likert-style scales on positive emotions and negative emotions, whereas some studies with categorical variables ask participants to compare the level of regret between action or inaction, by choosing one of the two options. All statistics were converted to Hedges’ *g* effects[[3]](#footnote-4), as this allowed for a generalized comparison of effect sizes between studies of different designs (see similar methods by Fillon et al., 2020; Jachimowicz et al., 2019).

Our meta-analysis included within-subject studies, between-subject studies, and one-sample comparison studies. We conducted a separate meta-analysis for comparison studies and for experimentally manipulated studies.

Table 4

*Designs, Variables, and Statistics of Included Studies*

|  |  |  |  |
| --- | --- | --- | --- |
| Type(s) of Variables Included | Features of Measurement with Example | Commonly Reported Statistics | Conversion of Statistics |
| Continuous Variable | For measuring the intensity of emotions, Likert scales, with different numbers of points across different studies  See Connolly et al. (1997) | 1) Descriptive statistics: *M*, *SD*.  2) Inferential statistics: t-statistics, F-statistics, and Cohen’s *d* | 1) For between-subject studies, if the original studies only provided *M* and *SD*, we calculated the Cohen’s *ds* with escalc function of metafor (Viechtbauer, 2010). For within-subject studies, if the original studies only reported *M* and *SD*, we calculated Cohen’s *dav* using d.dep.t.avg function of [MOTE v1.0.2](https://www.rdocumentation.org/packages/MOTE/versions/1.0.2) (Buchanan et al., 2019). According to Lakens (2013), *dav* is more similar (relative to *drm*) to *ds* under most situations.  2) Between-subject t-statistics were converted into Cohen’s *ds* with esc\_t function of esc v0.5.1 (Lüdecke, 2019), whereas within-Subject t-statistics is converted into Cohen’s *dz*with Lakens (2013) Formula 7: t-statistics/sqrt(n). That said, as Lakens (2013) mentioned, *dz* is less preferred in meta-analyses especially when researchers need to compare effects between within-subject studies and between-subject studies. We planned to obtain *M* and *SD* to calculate *dav* (more similar to *ds*), a suggested solution by Lakens (2013), from authors for within-subject studies if t-statistics is provided in the article but *M* and *SD* are not provided. If we couldn’t obtain *M* and *SD*, we would calculate *dz*, with t-statistics and sample size[[4]](#footnote-5).  3) Then we converted Cohen’s d into Hedges’ g with d\_to\_g function of MAc v1.1 (Re, 2012) |
| Categorical Variable | Binary Choice, comparing action and inaction  See Kahneman and Tversky (1982) example | 1) Descriptive statistics: % or count of higher emotional intensity for action and inaction  2) Inferential statistics: Chi-Square statistics, Cramer *V*, etc. | 1) If Chi-square was not provided in the original article, then we converted counts of action-inaction emotion into Chi-square statistics with chisq.test, a base R function  2) Chi-square is Converted to Cohen’s d with chies function of compute.es v0.2-5 (Re, 2020).  3) Then Cohen’s *d* was converted to Hedges’ *g* with d\_to\_g function of [MAc v1.1](https://www.rdocumentation.org/packages/MAc/versions/1.1) (Re, 2012) |

We planned to include studies with three types of dependent variables - positive emotions, negative emotions, and counterfactual thought. The dichotomous categorization between positive emotions and negative emotions may not always be clear and may be overly simplistic as some emotions may reflect a mixture of positivity and negativity (An et al., 2017). Our classification is based on pleasantness, as positive emotion, versus unpleasantness, as negative emotion (Schlosberg, 1954), but not based on other dimensions such as whether the emotion is functional or not. Here, we specify that in the context of action effect, positive emotions include measures notably but not limited to happiness, satisfaction, joy, elation, feeling good, relief, rejoicing, whereas negative emotions include measures such as but not limited to regret, disappointment, guilt, sadness, feeling bad, etc. For positive emotion, for example, in Connolly et al. (1997), the experimenters presented participants with the question “*How happy overall do you think each student would be the course section he ended up in?*” (p. 76). Participants rated the level of happiness on an 11-point scale, “from -5 (*very unhappy*) to +5 (*very happy*)” (p. 76). For negative emotion, for example, in Zeelenberg et al. (2002) Experiment 2, the experimenters presented participants with this question: “*How much would Steenland [Straathof] regret his decision?*”, on a 7-point regret scale (“1 = *very little*, 7 = *very much*”) (p. 317). For categorical variables, for example, Kahneman and Tversky (1982) asked participants “Who feels more regret?”, comparing Paul who did not switch vs George who switched. In both cases, we coded the valence of the emotion.

## Confirmatory Analyses

We used RStudio 1.3.1093 with RMarkdown (RStudio Team, 2021; Xie et al., 2018) and metafor 2.4-0 version (Viechtbauer, 2010) for the statistical analyses. For main-effects, we analyzed with separate two-level random-effects models and three-level models (Cheung, 2019), for experimentally manipulated studies and comparison studies. We adopted random-effects models as we expected there would be substantial heterogeneity (Slaney et al., 2018), given the mixed and opposite findings in action-effect/inaction-effect. Furthermore, we used three-level models (with metafor, Viechtbauer, 2010) to account for dependencies of effect sizes within the same article (Cheung, 2019). Since some moderators (e.g., hypothetical vs real life experience and temporal distance) may be associated with each other, we conducted chi-square tests for the associations between moderators (referring to Hofmann et al., 2010; Lipsey, 2003 Table 2). This is reported in Supplementary Table 7. We reported the results of three-level models in the main manuscript as they are more accurate, and results of two-level models in the supplementary. For moderator analyses, we used both meta-regression models, two-level fixed-effects models, and three-level models for contrasting moderator categories, as well as MetaForest (van Lissa, 2017).

We converted all effect sizes into Hedge's *g* to facilitate comparison. If the original study consisted of more than one action condition and/or more than one inaction condition (e.g., studies with more than one factor), we collapsed them into two conditions, given sufficient data. If a study consists of multiple scenarios with multiple effect sizes, we combined them into one effect size in our main-effect analyses. Apart from conducting three-level analyses for both main effects and moderator analyses, as exploratory robustness checks we added agg function (MAd package, Re & Hoyt, 2014) for effect aggregation of two-level models as there might be a few studies with multiple dependent effect sizes. In general, assuming the correlation between the measures to be 0.5 (as most studies did not report correlations, Wampold et al., 1997) is the common practice but we also conducted sensitivity analyses assuming correlations to be 0.3 and 0.7. We reported results based on agg function in the supplementary Table 8 and reported results based on three-level model in the main manuscript. As suggested by the reviewer Dan Quintana, the aggregation method may result in losses of precision, so we prefer using the three-level model to account for effect size dependencies. That said, the two-level model results using agg function, assuming correlations to be 0.3, 0.5, or 0.7, are very similar to results using three-level model. If a study consists of multiple dependent variables, each dependent variable with its effect size and variance were incorporated into the main effect analyses.

Whenever available, we collected descriptive statistics (e.g., mean, *SD*) or inferential statistics (e.g., t-statistics), directly from authors of original papers or the original article to calculate the effect sizes. We checked for the accuracy of these analyses based on the provided information and details. In most cases if *M, SD* and sample sizesare given, for between-subject studies, we analyzed using *ds*, whereas for within-subject studies, we analyzed using *dav*as they are more comparable to *ds* (see our Table 4, Lakens, 2013).However, if *M*, *SD*, and sample sizes are not given, we either calculate *dz* based on the t-statistics and the sample size, or we use the reported Cohen’s d from the original study. These are not ideal solutions, but these solutions are likely reasonable estimates given insufficient information. In cases with large discrepancies between our calculated effect size and reported effect size, we first double-checked if our conversion method is correct, and then contacted the original authors for further information, as it is possible the original authors made errors in their original calculations or reports.

Our decisions were clearly documented, based on information available in the original article or obtained from the original authors. We believe it is unlikely that such discrepancies in very few studies would result in substantial differences in synthesis results, but if that is the case, we disclose them transparently. Whenever standardized effect sizes were not unavailable, we used either descriptive statistics or inferential statistics, such as Mean and Standard Deviation, Chi-Square Statistics, Count, and t-statistics to re-compute standardized effect size, using several R packages, e.g., powerAnalysis (Fan, 2017), esc (Lüdecke, 2019), effsize (Torchiano, 2020). We documented all conversions and coding decisions and included the original quotes and/or page/table numbers from the original articles into the coding sheet to facilitate reproducibility.

There are many approaches to publication bias methods, and the suitability of different methods depends on factors such as the number of studies and heterogeneity (Carter et al., 2019; Rodgers & Pustejovsky, 2020; van Aert et al., 2016), on which we have limited information at Stage 1 before data extraction. We stated our planned preferred effect size adjustment methods under different scenarios in Supplementary Table 5, based on Carter et al. (2019) simulations of false-positive rate and statistical power.

## Priori Power Analysis

Since our understanding of the literature suggests that the action-effect is robust, we expected the effect size of the action-effect to be *g* = 0.5, 95% CI [0.20, 0.80]. We aimed to detect the lower confidence interval of this effect (g = 0.2). We expected to include 20-50 samples for comparison studies and 20-50 samples for experimental studies. We expected the average sample size per study to be 80-200 because earlier studies tend to have smaller sample sizes (under 100, e.g., Gilovich & Medvec, 1994) but the more recent well-powered studies (sample of over 300 participants, e.g., Feldman, 2020) may boost up the average sample size. We expected heterogeneity to be high, due to some opposite direction findings within the literature. We conducted a priori-power calculation based on the lower confidence intervals and lower limits of our estimates above (*k* = 20, sample size = 80, *g* = 0.2) with code developed by Vembye et al. (2022) for multi-level meta-analyses. The expected estimated power is 76.3% given the above estimations. We expect statistical power for both comparison studies and experimental studies (assuming the effect sizes of comparison studies and experimental studies are both 0.2 and above, with 20 or more samples for both types of studies) to be adequate. We note that such estimations are conservative, and the power is likely higher. For example, the statistical power would be >99.9% if we assume *k* = 35, sample size = 140, *g* = 0.5, a more optimistic estimation.

We planned to focus our reporting of moderator analyses on three-level model, yet we will also conduct fixed-effects two-level model analyses and MetaForest analyses.

# Results

[Note 1: The below results are based on simulated data, and will be replaced by real results in Stage 2 after search/coding]

[Note 2: As suggested by the reviewer Dan Quintana, we intend to conduct Bayesian Analyses in cases of disagreements of publication bias tests mentioned above. This means that for experimental and/or comparison studies, if the agreement rate between adjustment tests falls below 5/6, we will conduct Bayesian Analyses. For example, a substantial disagreement occurs when four tests find support for an effect with confidence intervals not overlapping with null whereas two tests fail to find support for an effect. In such scenarios, we will conduct Bayesian analyses. We expect this to be very unlikely. Based on our knowledge of the literature, even though action-effect is weakened in some conditions and reversed in a few conditions, action-effect is generally a robust and replicable effect (see the partial list of successful replications in the Introduction Action-Effect section)]

We provided the list of studies included in the meta-analysis in Table 5. We summarized our findings in Table 6, moderator analysis in Table 8, and publication bias in Table 7. We presented the forest and funnel plots of the included samples in Figure 1 and Figure 2. In the following, we first present the main effect findings, followed by moderator analyses findings, and publication bias findings at the end.

Table 5

*Studies included in the Meta-analysis*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Study | N | Country | Sample population | Design | Publication status | DV type |
| Positive Emotions Experimental Studies | | | | | | | |
| 1. |  |  |  |  |  |  |  |
| 2. |  |  |  |  |  |  |  |
| 3. |  |  |  |  |  |  |  |
| 4. |  |  |  |  |  |  |  |
| Negative Emotions Experimental Studies | | | | | | | |
| 5. |  |  |  |  |  |  |  |
| 6. |  |  |  |  |  |  |  |
| 7. |  |  |  |  |  |  |  |
| 8. |  |  |  |  |  |  |  |
| Counterfactual Thoughts Experimental Studies | | | | | | | |
| 9. |  |  |  |  |  |  |  |
| 10. |  |  |  |  |  |  |  |
| 11. |  |  |  |  |  |  |  |
| 12. |  |  |  |  |  |  |  |
|  | | | | | | | |
| Positive Emotions Comparison Studies | | | | | | | |
| 13. |  |  |  |  |  |  |  |
| 14. |  |  |  |  |  |  |  |
| 15. |  |  |  |  |  |  |  |
| 16. |  |  |  |  |  |  |  |
| Negative Emotions Comparison Studies | | | | | | | |
| 17. |  |  |  |  |  |  |  |
| 18. |  |  |  |  |  |  |  |
| 19. |  |  |  |  |  |  |  |
| 20. |  |  |  |  |  |  |  |
| Counterfactual Thoughts Comparison Studies | | | | | | | |
| 21. |  |  |  |  |  |  |  |
| 22. |  |  |  |  |  |  |  |
| 23. |  |  |  |  |  |  |  |
| 24. |  |  |  |  |  |  |  |

Table 6

*Summarized Results of the Meta-Analysis*

|  |  |
| --- | --- |
| Hypotheses | Findings in the meta-analysis (Supported / Not Supported / Partially Supported) |
| *Main hypothesis* |  |
|  | (Supported / Not Supported / Partially Supported) |
| *Theoretical Moderator Hypotheses* |  |
|  | (Supported / Not Supported / Partially Supported) |

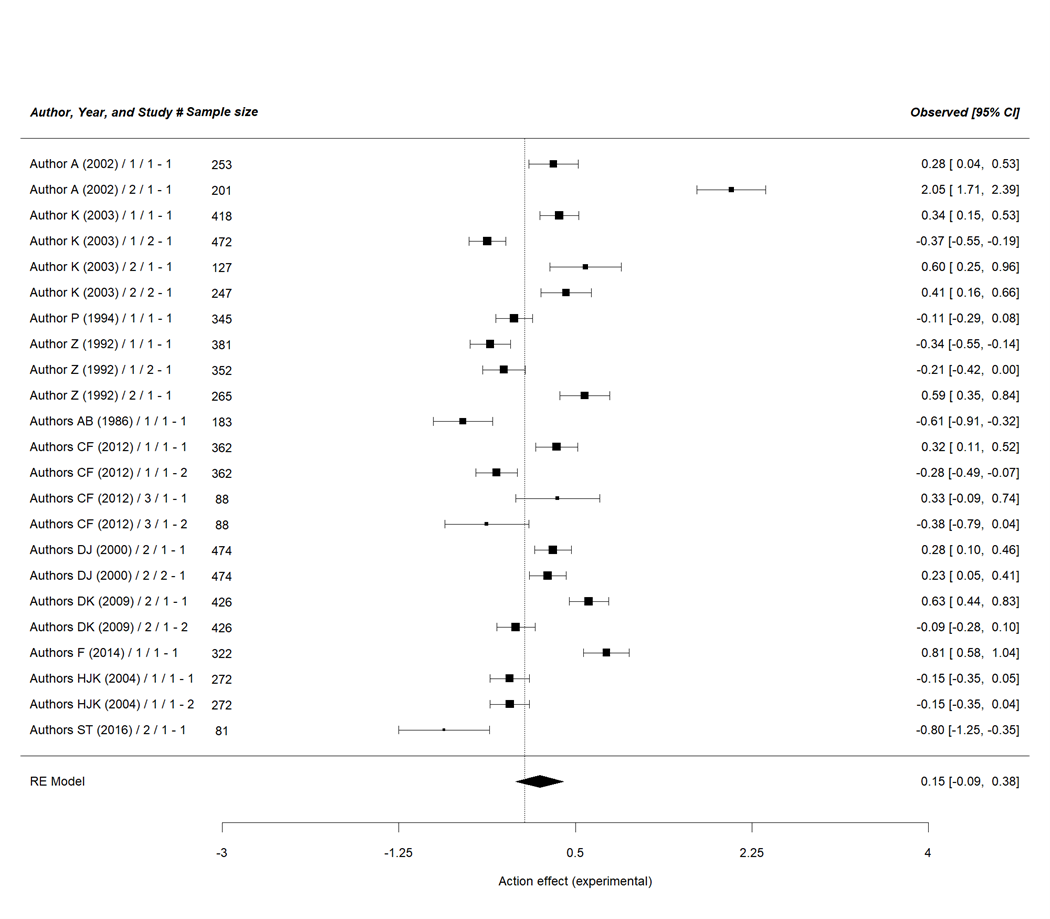
## Action-effect: Main effects

*Experimental Studies*

We first examined the overall effect of action-inaction on emotions and counterfactual thoughts for experimental studies through three-level models (see the Supplementary for two-level models results). Combining dependent variables, we found an overall effect of *g* = 0.15, CI [-0.09, 0.38]. The mean effect for positive emotions experimental studies was positive (*k* = 6, *g* = 0.42, CI [-0.41, 1.26]) whereas the mean effect for negative emotions experimental studies was negative (*k* = 12, g = -0.14, CI [-0.31, 0.03]). The mean effect for counterfactual thoughts was positive (*k* = 8, *g* = 0.28, CI [0.02, 0.55]). We failed to find support for the hypothesis for emotions, with the confidence intervals overlapping with null. This suggests that across the selected studies, we failed to find a meaningful difference in intensity of emotions between action and inaction. We failed to find support for a meaningful difference between positive and negative emotions in effect sizes. However, we found support for the hypothesis for counterfactual thoughts. See Figure 1 for the forest plot of all experimental studies. As suggested by the reviewer Dr. Emiel Cracco, we reported results of sensitivity analyses in Supplementary Table 9, excluding studies with *dz*as effect sizes. The number of studies after excluding *dz* studies is substantially lower and the effect sizes appear larger.

Figure 1

*Forest Plot for Experimental Studies*

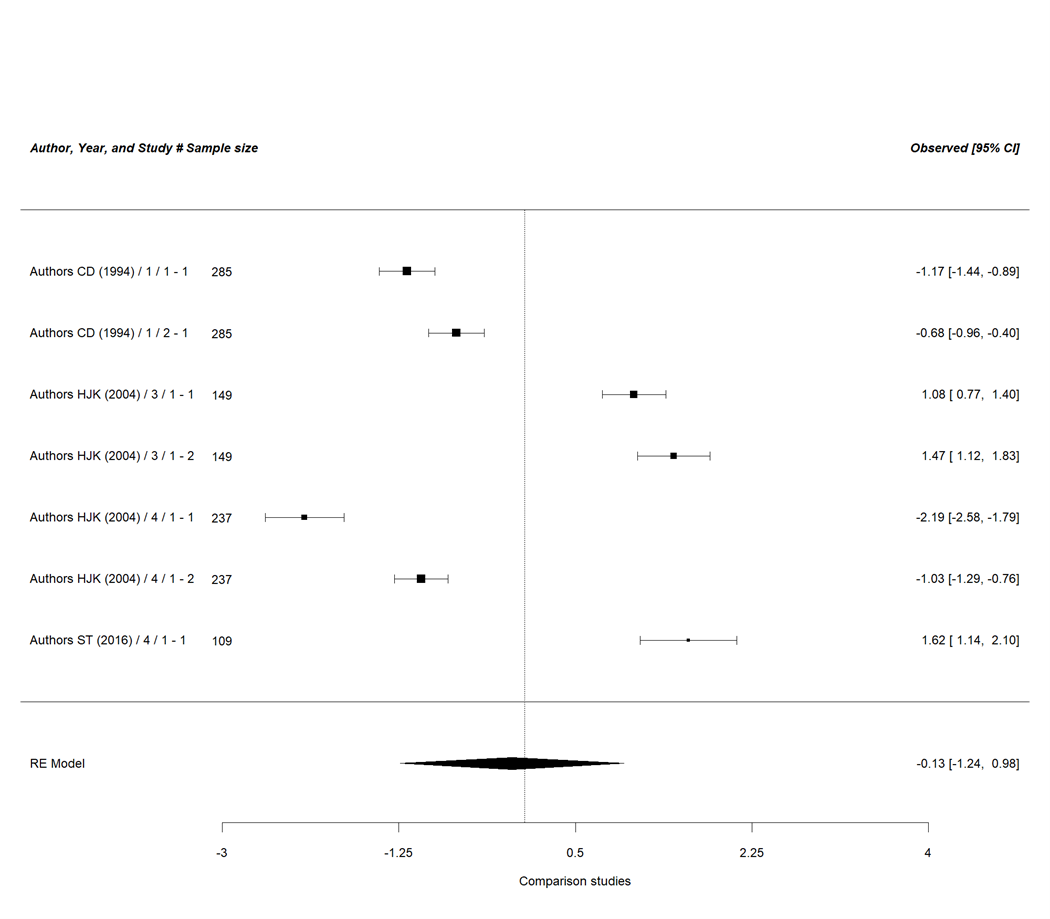


*Comparison Studies*

We first examined the overall effect of action-inaction on emotions and counterfactual thoughts for comparison studies through three-level models. Combining dependent variables, we found an overall effect of *g* = -0.12, CI [-1.21, 0.97]. The mean effect for positive emotions comparison studies was positive (*k* = 2, *g* = 0.21, CI [-1.55, 1.97]) whereas the mean effect for negative emotions comparison studies was negative (*k* = 4, *g* = 0.02, CI [-1.67, 1.70]). We failed to find support for a meaningful difference between positive and negative emotions. The mean effect for counterfactual thoughts was positive (*k* = 4, *g* = 0.03, CI [-1.64, 1.70]). We failed to find support for the hypothesis for emotions and counterfactual thoughts, with the confidence intervals overlapping with null. There is no difference in results before and after the exclusion of studies that only *dz* can be calculated, as comparison studies do not report t-statistics.

Figure 2

*Forest Plot for Comparison Studies*

**

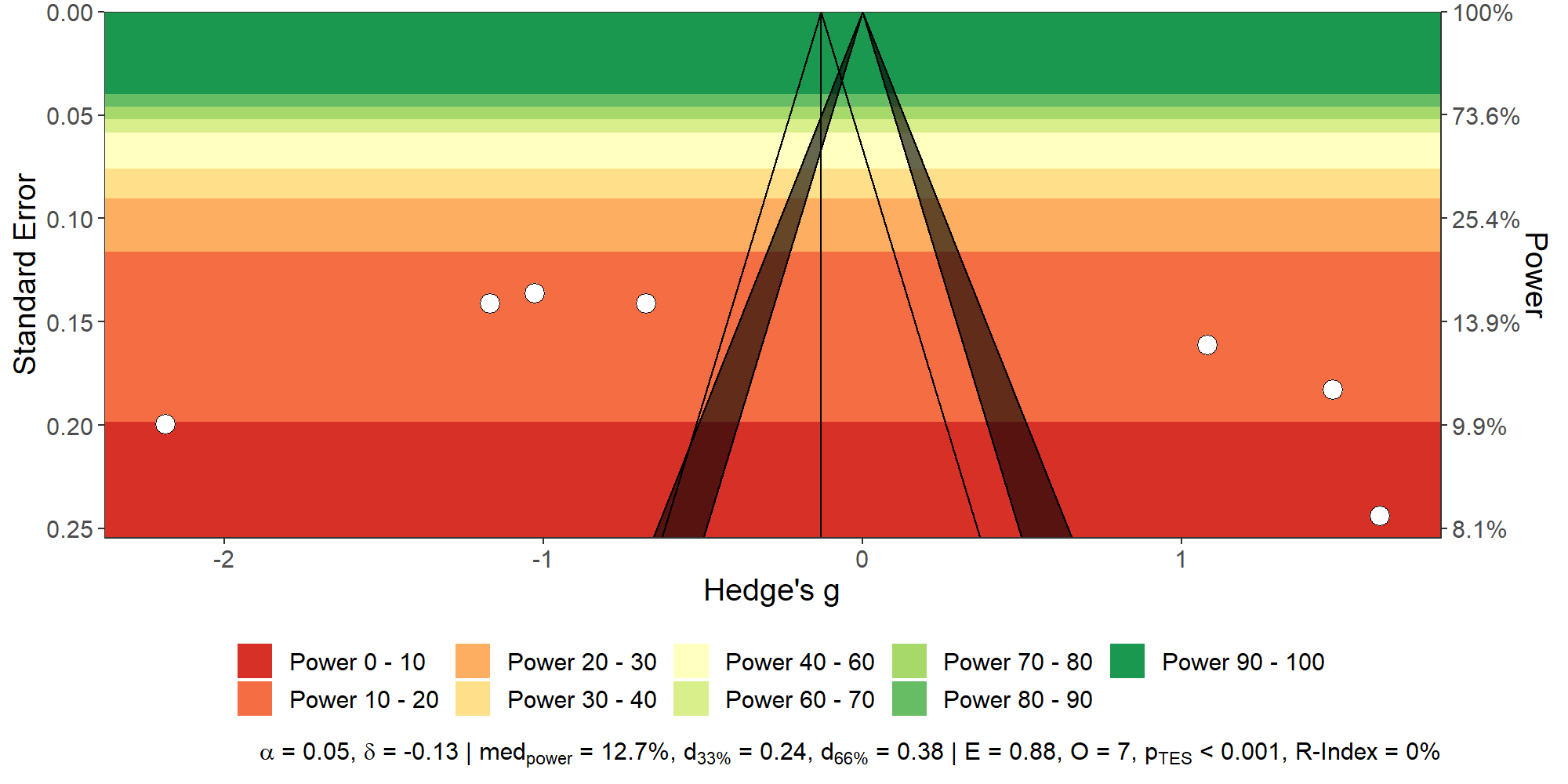
## Publication bias

Null findings are less likely to be published (Begg & Berlin, 1988; Carter et al., 2019), resulting in biased published literature and a possible overestimation of an effect. We employed nine different statistical approaches to examine a potential publication bias, including PET, PEESE (Stanley & Doucouliagos, 2014), P-Uniform (van Assen et al., 2015), Three-parameter Selection model (Iyengar & Greenhouse, 1988), Henmi and Copas (2010) confidence intervals, P-Curve (Simonsohn et al., 2014), as well as funnel plot asymmetry tests including Rank test (Begg & Mazumdar, 1994) and Egger’s Regression test (Sterne & Egger, 2005). See Figure 3 and Figure 4 for funnel plots (which assume *g* = -0.13 and *g* = 0.15 to be the true effect sizes for comparison studies and experimental studies respectively, as reminded by the reviewer Dan Quintana). Funnel plot asymmetry may also be due to other factors, such as low methodological quality in studies with smaller sample sizes and inflated effect sizes, artefactual causes or pure chance, not always due to publication bias (Egger et al., 1997; Page et al., 2021). We conducted several methods to compare if there are differences in results or interpretations of publication bias, and to assess publication bias more comprehensively. Some methods would be more suitable given the characteristics of our meta-analysis (factors such as the number of studies and heterogeneity, see Carter et al., 2019), which we discuss in the Discussion section. We provided more details regarding preferred methods under different scenarios (number of studies and heterogeneity), based on Carter et al. (2019) simulations, in Supplementary Table 5.

A summary of publication bias analyses is provided in Table 7. For comparison studies, the bias findings were not conclusive as we failed to detect bias with some methods but found support for presence of bias with some methods. This seems to be suggestive of a possible publication bias. With rank correlation tests and Egger’s regression tests, which are based on funnel plot asymmetry (see Figure 3 for funnel sunset plot), we failed to find support for evidence of publication bias. With corrections through p-uniform and p-curve, the effects are meaningful. But similarly, with Henmi and Copas (2010) method, Three-Parameter Selection Model, and PET-PEESE, the confidence intervals overlap with the null. Three-Parameter Selection Model is likely the most appropriate method given that the heterogeneity is high and there are limited comparison studies. We recognize that the median power of comparison studies is 12.7% so we restrain from drawing an accurate conclusion regarding the effect size magnitude. For experimental studies, the bias findings are mixed and seem to suggest the presence of publication bias. There are discrepancies using different publication bias methods. With rank correlation tests and Egger’s regression tests, which are based on funnel plot asymmetry (see Figure 4 for funnel sunset plot), we failed to find support for evidence of publication bias. With corrections using p-uniform and p-curve, the effects are meaningful. However, with other methods (Henmi & Copas, 2010, Three-Parameter Selection Model, and PET-PEESE), the confidence intervals overlap with the null. PET-PEESE is likely the most appropriate method given high heterogeneity and there are just under 30 studies. We note that the median power of experimental studies is 28.5% so we cannot draw a strong conclusion regarding the effect size magnitude. After excluding *dz* studies, most of the findings are similar. Similarly, we failed to find support for funnel plot asymmetry, whereas effect size confidence intervals overlap with null when we adopt Henmi and Copas (2010) test, Three-Parameter Selection Model, and PET-PEESE. We found support for larger effects with P-Uniform after exclusion compared to pre-exclusion but smaller effects with P-Curve after exclusion compared to pre-exclusion. Please see Supplementary Table 11 for the statistics.

Figure 3

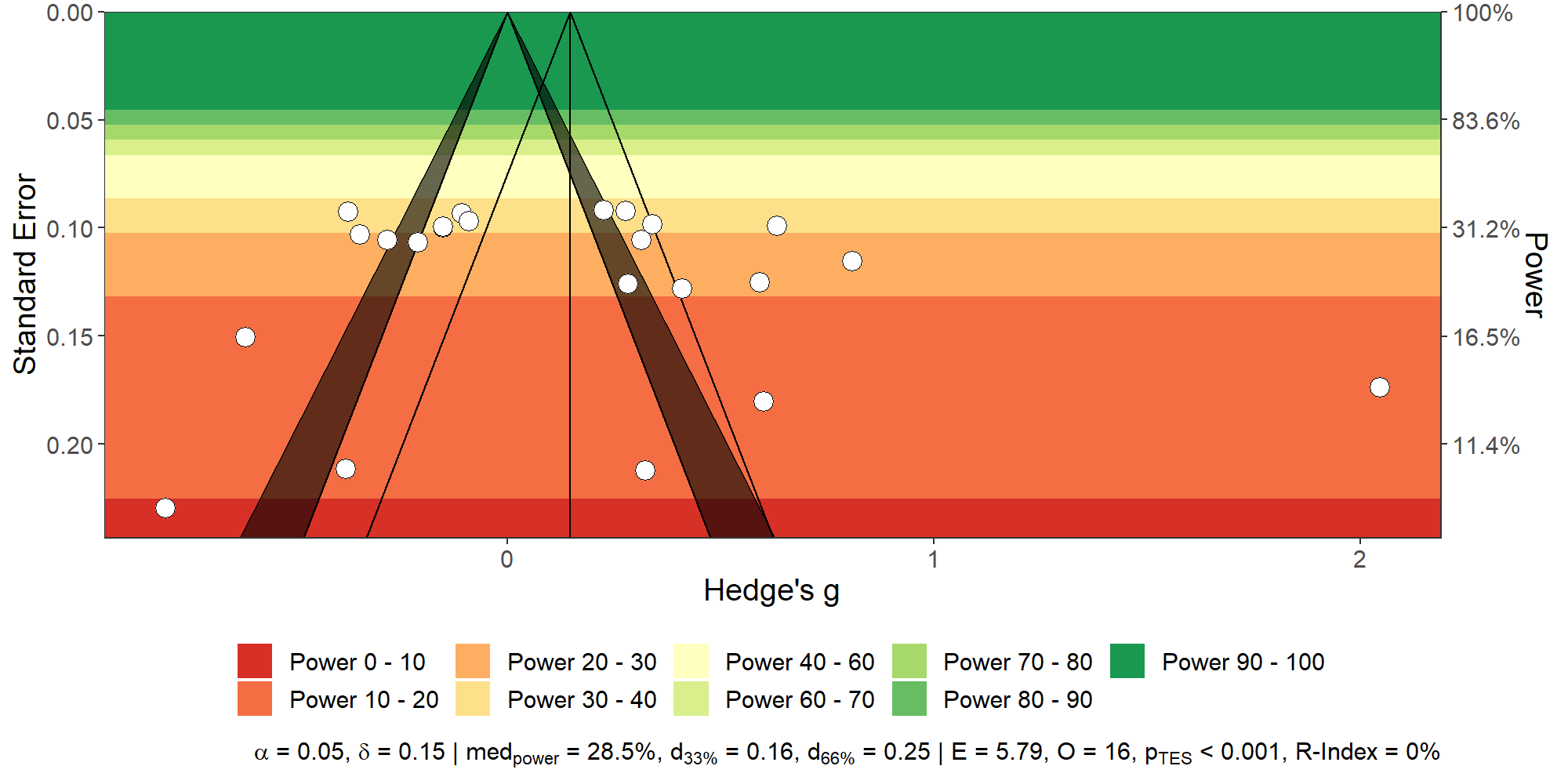
*Funnel Plot for Comparison Studies*



*Note.* We created the sunset plot with metaviz (Kossmeier et al., 2020), in which different colors represent different ranges of statistical power. Note that these power analyses assume that *g* = -0.13 is the true effect size.

Figure 4

*Funnel Plot for Experimental Studies*



*Note.* The above power analyses are based on the assumption that *g* = 0.15 is the true effect size.

Table 7

*Publication bias analyses*

|  |  |
| --- | --- |
| Publication bias analysis method | Results and adjusted models |
| Three-parameter selection model | Experimental studies:  Likelihood Ratio Test: 0.62, *p* = .400 Adjusted Model: *g* = -0.03, 95% CI [-0.34, 0.40]  Comparison studies:  Likelihood Ratio Test: 3.12, p = .600  Adjusted Model: g = -0.50, 95% CI [-2.04, 1.05] |
| PET | Experimental studies:  *b* = -0.18 [-1.02, 0.66]  Comparison studies:  b = -3.16 [-9.97, 3.65] |
| PEESE | Experimental studies:  *b* = 0.03 [-0.36, 0.41]  Comparison studies:  b = -1.58 [-4.89, 1.72] |
| P-Uniform | Experimental studies:  Adjusted Model: *g =* 0.11, 95% CI [0.07, 0.16], 11 studies significant  Comparison studies:  Adjusted Model: *g =* -0.34, 95% CI [-0.46, -0.22], 3 studies significant |
| Henmi and Copas (2010) | Experimental studies:  Adjusted Model: *g* = 0.21, 95% CI [-0.11, 0.33]  Comparison studies:  Adjusted Model: *g* = -0.34, 95% CI [-1.65, 0.97] |

|  |  |
| --- | --- |
| Publication bias analysis method | Results and adjusted models |
| Rank correlation test  (Begg & Mazumdar, 1994) | Experimental studies:  Kendall's tau = .06, *p* = .715  Comparison studies:  Kendall's tau = .20, *p* = .543 |
| Egger's regression test | Experimental studies:  *z* = 0.14, *p* = .891  Comparison studies:  *z* = 1.08, *p* = .280 |
| P-Curve | Experimental studies:  Evidential value: Yes; P-Curve Adjusted Cohen’s d = 0.42  Comparison studies:  Evidential value: Yes; P-Curve Adjusted Cohen’s d = 1.06 |

*Note*. 1) Values in parentheses indicate 95% confidence intervals [lower bound, upper bound], 2) Check Action-effect-meta-syntax-markdown-v2-K.html for their R outputs, 3) We will highlight and clearly specify the most relevant approach given information available by Stage 2 and elaborate further in the discussion section.

## Moderator analyses

Statistical heterogeneity between studies was determined using Cochran’s

Q statistics (Higgins & Thompson, 2002), which informs us about the presence or absence of heterogeneity (Huedo-Medina et al., 2006) and quantified using *I2* (Higgins & Thompson, 2002), which calculates the percentage of variance in effect due to real heterogeneity (Huedo-Medina et al., 2006). This global meta-analysis yielded a point estimate, confidence interval, and *p*-value, along with statistics for heterogeneity, assessed using Q-statistics and I-square. If there was indeed meaningful heterogeneity, we investigated and explored potential moderators.

The heterogeneity of effects was high. We ran a Cochran’s *Q* test and found support for heterogeneity in effect sizes across studies (*Q* (27) = 434.15, *p* < .001; *I2* = 95.8%), indicative of high variance (Higgins et al., 2003). The variation in effect-sizes was greater than would be expected from sampling error alone, indicating that moderator variables might be accountable for the variance in the effects.

We examined five possible theoretical and methodological moderators according to the pre-registered criteria and coding sheet: temporal distance, study design, valence or information of prior outcome, specific vs general, and hypothetical vs real-life events. We summarized the results of the moderator analyses in Table 8.

In the following, we reported results of different models. We note that when the total number of studies is under 30, we prefer interpreting the results based on MetaForest, as it can account for limited statistical power given a limited number of studies (van Lissa, 2020). If the total number of studies is over 30, we prefer three-level models as they can take into account the possible dependence of effect sizes with the same article (Cheung, 2019). We set this threshold arbitrarily, as no study has compared performances between MetaForest, traditional mixed-effects two-level model, and three-level model given different numbers of studies. We would appreciate constructive feedback from reviewers.

We report findings of sensitivity analyses (excluding *dz* effects) in Supplementary Table 10, in which there are some discrepancies in findings. We briefly describe these findings below.

### Temporal Distance.

We adopted metaregressions, with temporal distance as a continuous variable. With the three-level metaregression model, we failed to find support, *β* = 0.01, 95% CI [-0.01, 0.03], *p* = .607. For the sensitivity analysis, we failed to find conclusive evidence for moderation, *β* = 0.04, 95% CI [-0.00, 0.08], *p* = .084. Additionally, we also conducted analyses with temporal distance as a categorical variable. Fifteen studies with current or events with lower temporal distance (within a year) had a mean effect of *g* = -0.26, CI [-0.65, 0.14] but confidence intervals overlap with null. Eleven studies with events that are more than 1 year ago but not lifetime had a mean effect of *g* = 0.32, CI [-0.06, 0.70] but confidence intervals overlap with null. Seven studies measuring emotions or counterfactual thoughts for major lifetime events had a mean effect of *g* = 0.46, CI [-0.31, 1.24]. With the three-level model, we found support for a moderating effect of temporal distance. Studies with more distant events had a larger effect. For the sensitivity analysis, we also found support for moderation of temporal distance.

### Study Design.

Ten studies with between-subject design had a mean effect of *g* = 0.40, CI [-0.05, 0.86]. Seven studies with one-sample comparison design had a negative mean effect, *g* = -0.13, CI [-1.24, 0.98]. Eleven studies with within-subject design had a mean effect of *g* = -0.03, CI [-0.25, 0.18]. With the random-effects three-level model, we failed to find support for a moderating effect of study design. For the sensitivity analysis, we also found no evidence for moderation.

### Valence or prior outcome information.

Seventeen studies with positive, neutral or no prior outcome information had a mean effect of *g* = -0.09, CI [-0.53, 0.35]. Seventeen studies with negative outcome had a mean effect of *g* = 0.22, CI [-0.16, 0.60]. We found support for a difference with the three-level model. We detected a meaningfully larger effect for studies with negative outcomes compared to studies with non-negative outcomes. This is in the opposite direction of our hypothesis. For the sensitivity analysis, we also found larger effects for studies with negative prior outcomes compared to other studies.

### Specific vs general.

Twenty-one studies for specific events/memories had a mean effect of *g* = 0.16, CI [-0.16, 0.48]. Thirteen studies for general events/memories had a minimal mean effect of *g* = 0.01, CI [-0.54, 0.56]. We found support for a moderation effect with the three-level model, with a stronger effect for specific events/memories than general events/memories. For the sensitivity analysis, we found larger effects for general events/memories, compared to specific events/memories.

### Hypothetical vs real-life events.

Eleven studies for hypothetical scenarios had a minimal mean effect of *g* = 0.05, CI [-0.51, 0.62]. Twenty-one studies for real-life events had a mean effect of *g* = 0.12, CI [-0.22, 0.47]. We failed to find support for the moderation effect with the random-effects three-level model. However, with the sensitivity analysis, we found that studies with hypothetical scenarios had larger effects than real-life events studies.

### Self vs Other.

Sixteen studies for rating self studies had a minimal mean effect of *g* = -0.10, CI [-0.48, 0.29]. Seventeen studies for rating other studies had a mean effect of *g* = 0.28, CI [-0.13, 0.69]. We failed to find support for the moderation effect with the random-effects three-level model. With the sensitivity analysis, we found that third-person studies had larger effects than self-rating studies.

### Meanings.

Thirteen studies with change or no change meanings had a minimal mean effect of *g* = 0.01, CI [-0.47, 0.48]. Nine studies with doing something vs not doing something/doing nothing meanings had a mean effect of *g* = 0.19, CI [-0.39, 0.77]. We found support with the random-effects three-level model. With the sensitivity analysis, we failed to find support for a moderating effect.

### Normality.

Twelve studies with action norm had a mean effect of *g* = 0.10, CI [-0.48, 0.67]. Nine studies with inaction norm had a negative mean effect of *g* = -0.11, CI [-0.67, 0.45]. We failed to find support for the moderation effect with the mixed-effects two-level model but found support with the random-effects three-level model. With the sensitivity analysis, we found that studies with action norm had larger effects than studies with inaction norm.

### MetaForest moderator analyses

To address the problem of limited studies and lack of statistical power for moderator analyses without risk overfitting, we adopted MetaForest (van Lissa, 2017). MetaForest uses "random forests", a kind of machine learning technique, and bootstrapping to examine several possible moderators. The main model indicator, R-squared (R-OOB) was -0.02. The negative value means that the model is overfitting with inclusion of some noise predictors (van Lissa, 2017). A positive variable importance value implies that the variable is a meaningful moderator whereas a negative variable importance value implies the variable is not a meaningful moderator. Temporal Distance is the most important variable (the highest in variable importance value, stronger effect for more distant events), followed by design type (stronger effect for between-subject studies, followed by within-subject studies, and comparison studies) and meanings (stronger effect for other or not specified studies, followed by something vs nothing studies, and change or no change studies). Real-life vs lab, self vs other, prior outcome, normality, and specific vs general have negative variable importance values. This means that these variables are not meaningful moderators. For the sensitivity analysis (excluding *dz* studies), the R-squared appears larger (0.10). The variables with negative importance values remain the same, apart from normality with a positive variable important value in the sensitivity analysis. Temporal distance remains the most important variable, followed by normality, design type, and meanings.

Table 8

*Moderators: Summarized results*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Moderator | *k* | *Q* | *df* | *g* | ‎95% CI‎ | Difference | *ML p* | |
| **Temporal Distance** |  |  |  |  |  |  |  | |
| Current/Short-Term (< 1 year) | 15 | 374.54 | 14 | -0.26 | -0.65, 0.14 |  |  | |
| Medium Term (1 year or more) | 11 | 137.30 | 10 | 0.32 | -.0.06, 0.70 | 0.57 | .004\*\* | |
| Lifetime | 7 | 263.30 | 6 | 0.46 | -0.31, 1.24 | 0.72 | .213 | |
| **Study Design** |  |  |  |  |  |  |  | |
| Between-Subject Design | 10 | 194.52 | 9 | 0.40 | -0.05, 0.46 |  |  | |
| Comparison Design | 7 | 392.17 | 6 | -0.13 | -1.24, 0.98 | 0.53 | .383 | |
| Within-Subject Design | 11 | 96.84 | 10 | -0.03 | -0.25, 0.18 | 0.40 | .088 | |
| **Prior Outcome** |  |  |  |  |  |  |  | |
| Negative | 17 | 450.72 | 16 | 0.22 | -0.16, 0.60 |  |  | |
| Neutral / Positive / No Information | 17 | 441.21 | 16 | -0.09 | -0.53, 0.35 | 0.30 | .019\* | |
| **Specific vs General** |  |  |  |  |  |  |  | |
| Specific | 21 | 480.52 | 20 | 0.16 | -0.16, 0.48 |  |  | |
| General | 13 | 367.12 | 12 | 0.01 | -0.54, 0.56 | 0.15 | .002\*\* | |
| **Hypothetical vs Real-life** |  |  |  |  |  |  |  | |
| Hypothetical | 11 | 354.59 | 10 | 0.05 | -0.51, 0/62 |  |  | |
| Real-life | 21 | 478.37 | 20 | 0.12 | -0.22, 0.47 | -0.07 | .117 | |
| **Meanings** |  |  |  |  |  |  |  | |
| Change vs No Change | 13 | 395.58 | 12 | 0.01 | -0.47, 0.48 |  |  | |
| Doing Something vs Doing Nothing/Not Doing Something | 9 | 234.15 | 8 | 0.19 | -0.39, 0.77 | -0.18 | <.001\*\*\* |
| **Normality** |  |  |  |  |  |  |  | |
| Action norm | 12 | 405.88 | 11 | 0.10 | -0.48, 0.67 |  |  | |
| Inaction norm | 9 | 283.67 | 8 | -0.11 | -0.67, 0.45 | 0.21 | <.001\*\*\* | |

*Note.* *k* = number of samples; *Q* = Q-statistics, *df* = degree of freedom, *g* = Hedge’s g effect size for studies of each moderator category, CI = lower and upper limits of 95% confidence interval, Difference = differences of effect sizes based on fixed-effect model, \* *p* < .05, \*\* *p* < .01, \*\*\* *p* <.001, based on Multilevel (*ML*) model (all two-tailed)

# Discussion

[TBD pending, to be completed after real data collection and data analysis]

## Limitations and future directions

[We will address the reviewer suggested future direction of addressing emotions other than regret and joy in action-inaction asymmetries.]

[If there are meaningful differences in findings before and after excluding studies in which we are only able to calculate *dz*, we will address and discuss the issue. In any case, we will emphasize that future studies should report both *M* and *SD*, with effect sizes, so that more accurate effect size estimates can be obtained]

## 

# References

Allen, L., & O’Connell, A. (2014). CRediT - Contributor Roles Taxonomy. Retrieved May 13, 2020, from <https://casrai.org/credit/>

An, S., Ji, L. J., Marks, M., & Zhang, Z. (2017). Two sides of emotion: exploring positivity and negativity in six basic emotions across cultures. *Frontiers in Psychology, 8*, 610. <https://doi.org/10.3389/fpsyg.2017.00610>

Anderson, C. J. (2003). The psychology of doing nothing: forms of decision avoidance result from reason and emotion. *Psychological Bulletin, 129*(1), 139. https://doi.org/10.1037/0033-2909.129.1.139

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3. <https://doi.org/10.1037/amp0000191>

Arizona Software Inc. (2010). GraphClick 3.0.2. http://www.arizona-software.ch/graphclick/

Arslan, R. C. (2019). How to automatically document data with the codebook package to facilitate data reuse. *Advances in Methods and Practices in Psychological Science, 2*(2), 169-187. <https://doi.org/10.1177/2515245919838783>

Bar-Eli, M., Azar, O. H., Ritov, I., Keidar-Levin, Y., & Schein, G. (2007). Action bias among elite soccer goalkeepers: The case of penalty kicks. *Journal of Economic Psychology, 28*(5), 606-621. <https://doi.org/10.1016/j.joep.2006.12.001>

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](https://doi.org/10.1037%2F1089-2680.5.4.323)

Begg, C. B., & Berlin, J. A. (1988). Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *151*(3), 419-445. https://doi.org/10.2307/2982993

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 1088-1101. <https://doi.org/10.2307/2533446>

Bonnefon, J. F., & Zhang, J. (2008). The intensity of recent and distant life regrets: An integrated model and a large scale survey. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(5), 653-662. https://doi.org/10.1002/acp.1386

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264. <https://doi.org/10.1037/hea0000294>

Buchanan, E. M., Gillenwaters, A., Scofield, J. E., & Valentine, K. D. (2019). MOTE: Measure of the Effect: Package to assist in effect size calculations and their confidence intervals. *R Package Version 1.02, 2019*

Byrne, R. M., & McEleney, A. (2000). Counterfactual thinking about actions and failures to act. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1318. https://doi.org/10.1037/0278-7393.26.5.1318

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science, 2*(2), 115-144. https://doi.org/10.1177/2515245919847196

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization, 81*(1), 1-8. <https://doi.org/10.1016/j.jebo.2011.08.009>

Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychology Review,* 1-10. <https://doi.org/10.1007/s11065-019-09415-6>

Connolly, T., Ordóñez, L. D., & Coughlan, R. (1997). Regret and responsibility in the evaluation of decision outcomes. *Organizational Behavior and Human Decision Processes, 70*(1), 73-85. <https://doi.org/10.1006/obhd.1997.2695>

Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science, 11*(6), 212-216. <https://doi.org/10.1111/1467-8721.00203>

Coricelli, G., & Rustichini, A. (2010). Counterfactual thinking and emotions: Regret and envy learning. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365* (1538), 241–247. https://doi.org/10.1098/rstb.2009.0159

Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C., & Thompson, S. C. (1995). The undoing of traumatic life events. *Personality and Social Psychology Bulletin, 21*, 109–124. https://doi.org/10.1177/0146167295212002

Davison, I. M., & Feeney, A. (2008). Regret as autobiographical memory. *Cognitive Psychology, 57*(4), 385-403. https://doi.org/10.1016/j.cogpsych.2008.03.001

DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science, 22*(4), 442-446. <https://doi.org/10.1177%2F0956797611400616>

Drevon, D., Fursa, S. R., & Malcolm, A. L. (2017). Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior modification, 41*(2), 323-339. <https://doi.org/10.1177/0145445516673998>

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629-634.

Fan, F. Y. (2017). Package ‘powerAnalysis’. *R Package Version 0.2.1, 2017*

Feldman, G. (2020). What is normal? Dimensions of action-inaction normality and their impact on regret in the action-effect. *Cognition and Emotion, 34*(4), 728-742. https://doi.org/10.1080/02699931.2019.1675598

Feldman, G., & Albarracín, D. (2017). Norm theory and the action-effect: The role of social norms in regret following action and inaction. *Journal of Experimental Social Psychology, 69*, 111-120. <https://doi.org/10.1016/j.jesp.2016.07.009>

Feldman, G., Chandrashekar, S. P., Yeung, S. K., Xiao, Q., Fillon, A., & Henne, P. (2021). *What is action, what is inaction? Clarifying action-inaction in judgment and decision making and recommendations for term use and typology.* [Manuscript in preparation].

Feldman, G., Kutscher, L., & Yay, T. (2020). Omission and commission in judgment and decision making: Linking action-inaction effects using the concept of normality. *Social and Personality Psychology Compass*. https://doi.org/10.1111/spc3.12557

Feldman, G., Wong, K. F. E., & Baumeister, R. F. (2016). Bad is freer than good: Positive–negative asymmetry in attributions of free will. *Consciousness and Cognition*, *42*, 26-40. <https://doi.org/10.1016/j.concog.2016.03.005>

Feldman, J., Miyamoto, J., & Loftus, E. F. (1999). Are actions regretted more than inactions? *Organizational Behavior and Human Decision Processes, 78*(3), 232-255. <https://doi.org/10.1006/obhd.1999.2833>

Feltz, A., & May, J. (2017). The means/side-effect distinction in moral cognition: A meta-analysis. *Cognition*, *166*, 314-327. <https://doi.org/10.1016/j.cognition.2017.05.027>

Fillon, A., Kutscher, L., & Feldman, G. (2020). Impact of past behavior normality on regret: Meta-analysis of exceptionality effect. *Cognition and Emotion, 35*(1), 129-149*.* <https://doi.org/10.1080/02699931.2020.1816910>

Fillon, A., 2022). Evaluations of action and inaction decision-makers in risky decisions resulting in negative outcomes: Inaction agents are preferred to and perceived as more competent and normative than action agents. https://doi.org/10.17605/OSF.IO/A8E4D

Flower, A., McKenna, J. W., & Upreti, G. (2016). Validity and reliability of GraphClick and DataThief III for data extraction. *Behavior Modification, 40*(3), 396-413. https://doi.org/10.1177/0145445515616105

Garrett, R. (1996). Three definitions of wisdom. In *Knowledge, teaching and wisdom* (pp. 221-232). Springer, Dordrecht.

Gilovich, T., & Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology, 67*(3), 357. <https://doi.org/10.1037/0022-3514.67.3.357>

Gilovich, T., & Medvec, V. H. (1995). The experience of regret: what, when, and why. *Psychological Review, 102*(2), 379. <https://doi.org/10.1037/0033-295x.102.2.379>

Gilovich, T., Medvec, V. H., & Kahneman, D. (1998). Varieties of regret: A debate and partial resolution. *Psychological Review, 105*(3), 602. https://doi.org/10.1037/0033-295X.105.3.602

Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A., & Sherman, S. J. (1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin, 16*(2), 284-295. https://doi.org/10.1177/0146167290162009

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology, 26*(2), 201-218. https://doi.org/10.1002/(SICI)1099-0992(199603)26:2<201::AID-EJSP745>3.0.CO;2-J

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta‐analysis and robustness to publication bias. *Statistics in Medicine*, *29*(29), 2969-2983. <https://doi.org/10.1002/sim.4029>

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta‐analysis. *Statistics in medicine*, *21*(11), 1539-1558. <https://doi.org/10.1002/sim.1186>

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560. <https://doi.org/10.1136/bmj.327.7414.557>

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological bulletin*, *136*(3), 390. [https://doi.org/10.1037/a0018916](https://psycnet.apa.org/doi/10.1037/a0018916)

Huang, W. H., & Zeelenberg, M. (2012). Investor regret: The role of expectation in comparing what is to what might have been. *Judgment & Decision Making, 7*(4).

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index?. *Psychological Methods, 11*(2), 193. <https://doi.org/10.1037/1082-989X.11.2.193>

Iyengar, S., & Greenhouse, J. B. (1988). Selection models and the file drawer problem. *Statistical Science*, 109-117. <https://doi.org/10.1214/ss/1177013012>

Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, *3*(2), 159-186. [s](https://doi.org/10.1017/bpp.2018.43)

Jamison, J., Yay, T., & Feldman, G. (2020). Action-inaction asymmetries in moral scenarios: Replication of the omission bias examining morality and blame with extensions linking to causality, intent, and regret. *Journal of Experimental Social Psychology, 89*, 103977. <https://doi.org/10.1111/spc3.12557>

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136. <https://doi.org/10.1037/0033-295X.93.2.136>

Kahneman, D. , & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47*, 263–291. https://doi.org/10.2307/1914185

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American, 246*(1), 160-173. <https://doi.org/10.1038/scientificamerican0182-160>

Kahneman, D., & Tversky, A. (1982b). On the study of statistical intuitions. *Cognition, 11*, 123–141. <https://doi.org/10.1016/0010-0277(82)90022-1>

Kossmeier, M., Tran, U. S., & Voracek, M. (2020). Power-enhanced funnel plots for meta-analysis. *Zeitschrift für Psychologie.* <https://doi.org/10.1027/2151-2604/a000392>

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology, 4*, 863. <https://doi.org/10.3389/fpsyg.2013.00863>

Lambdin, C., & Shaffer, V. A. (2009). Are within-subjects designs transparent? *Judgment and Decision Making, 4*(7), 554–566. https://doi.org/10.1037/e722352011-194

Landman, J. (1987). Regret and elation following action and inaction: Affective responses to positive versus negative outcomes. *Personality and Social Psychology Bulletin, 13*(4), 524-536. https://doi.org/10.1177/0146167287134009

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science, 587*(1), 69-81. https://doi.org/10.1177/0002716202250791

Lüdecke, D. (2019). Package ‘esc’. *R Package Version 0.5.1, 2019*

Moreau, D., & Gamble, B. (2020, January 7). Meta-analysis templates and materials. Retrieved from osf.io/q8stz

Mummolo, J., & Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review, 113*(2), 517-529. http://dx.doi.org/10.2139/ssrn.2956147

N'gbala, A., & Branscombe, N. R. (1997). When does action elicit more regret than inaction and is counterfactual mutation the mediator of this effect? *Journal of Experimental Social Psychology, 33*(3), 324-343. <https://doi.org/10.1006/jesp.1996.1322>

Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. *Advances in Methods and Practices in Psychological Science, 3*(2), 229–237. <https://doi.org/10.1177/2515245920918872>

Olsen, A. L. (2017). Responding to problems: actions are rewarded, regardless of the outcome. *Public Management Review, 19*(9), 1352-1364. <https://doi.org/10.1080/14719037.2017.1281998>

Page, M. J., Sterne, J. A., Higgins, J. P., & Egger, M. (2021). Investigating and dealing with publication bias and other reporting biases in meta‐analyses of health research: A review. *Research Synthesis Methods, 12*(2), 248-259.

Poisot, T., Sachse, R., Ashander, J., & Galili, T. (2016). Package “digitize”. *R Package Version 0.0.4, 2016*

Polanin, J. R., Pigott, T. D., Espelage, D. L., & Grotpeter, J. K. (2019). Best practice guidelines for Abstract screening large‐evidence systematic reviews and meta‐analyses. *Research Synthesis Methods, 10*(3), 330-342. <https://doi.org/10.1002/jrsm.1354>

Rajagopal, P., Raju, S., & Unnava, H. R. (2006). Differences in the cognitive accessibility of action and inaction regrets. *Journal of Experimental Social Psychology, 42*(3), 302-313. https://doi.org/10.1016/j.jesp.2005.05.003

Re, A. C. D. (2012). Package “MAc”. *R Package Version 1.1, 2012*

Re, A. C. D. (2020). Package “compute.es”. *R Package Version 0.2-5, 2020*

Re, A. C. D., & Hoyt, W. T. (2014). Package “MAd”: Meta-Analysis with Mean Differences. *R Package Version 0.8-2, 2014*

Reb, J., & Connolly, T. (2010). The effects of action, normality, and decision carefulness on anticipated regret: Evidence for a broad mediating role of decision justifiability. *Cognition and Emotion, 24*(8), 1405-1420. https://doi.org/10.1080/02699930903512168

Rodgers, M. A., & Pustejovsky, J. E. (2020). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods.* https://doi.org/10.1037/met0000300

RStudio Team. (2021). Integrated development for R. *RStudio, IncBoston, MA*.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61*(2), 81. <https://doi.org/10.1037/h0054570>

Sepehrinia, M., Baniyaghoub, R., & Nejat, P. (2022). *The Impact of Actor's Initial Status of Engagement in a Course of Action on Judgments of Post-Decisional Regret and Joy.* [Manuscript submitted for publication].

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*, 747-770. <https://doi.org/10.1146/annurev-psych-010418-102803>

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123-1128. https://doi.org/10.1177/1745691617708630

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9*(6), 666-681. <https://doi.org/10.1177/1745691614553988>

Slaney, K. L., Tafreshi, D., & Hohn, R. (2018). Random or fixed? An empirical examination of meta-analysis model choices. *Review of General Psychology, 22*(3), 290-304. https://doi.org/10.1037/gpr0000140

Stanley, T. D., & Doucouliagos, H. (2014). Meta‐regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60-78. <https://doi.org/10.1002/jrsm.1095>

Sterne, J. A., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, 99-110. <https://doi.org/10.1002/0470870168.ch6>

Torchiano, M. (2020). Package ‘effsize’. *R Package Version 0.8.0, 2020*

Towers, A., Williams, M. N., Hill, S. R., Philipp, M. C., & Flett, R. (2016). What makes for the most intense regrets? Comparing the effects of several theoretical predictors of regret intensity. *Frontiers in psychology, 7*, 1941. <https://doi.org/10.3389/fpsyg.2016.01941>

Tummers, B. (2006). DataThief III v.1.1. Available from http://www.datathief.org/

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Conducting meta-analyses based on p-values: reservations and recommendations for applying p- uniform and p-curve. *Perspectives on Psychological Science. 11*, 713–729. https://doi.org/ 10.1177/1745691616650874

van Assen, M. A., van Aert, R., & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods, 20*(3), 293. https://doi.org/10.1037/met0000025

van Lissa, C. J. (2017). MetaForest: Exploring heterogeneity in meta-analysis using random forests. Retrieved from <https://psyarxiv.com/myg6s/>

van Lissa, C. J. (2020). Small sample meta-analyses: Exploring heterogeneity using MetaForest. In *Small Sample Size Solutions* (pp. 186-202). Routledge.

Vembye, M. H., Pustejovsky, J. E., & Pigott, T. (2022). Power Approximations for Meta-Analysis of Dependent Effect Sizes. https://osf.io/6tp9y/download

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1-48. <https://doi.org/10.18637/jss.v036.i03>

Walters, W. H. (2007). Google Scholar coverage of a multidisciplinary field. *Information Processing & Management*, *43*, 1121-1132. <https://doi.org/10.1016/j.ipm.2006.08.006>

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A metaanalysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin, 122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide*. CRC Press.

Yeung, S. K., & Feldman, G. (2022). Revisiting the Temporal Pattern of Regret in Action Versus Inaction: Replication of Gilovich and Medvec (1994) With Extensions Examining Responsibility. *Collabra: Psychology, 8*(1): 37122. https://doi.org/10.1525/collabra.37122

Yeung, S. K., Yay, T., & Feldman, G. (2021). Action and inaction in moral judgments and decisions: ‎Meta-analysis of Omission-Bias omission-commission asymmetries. *Personality and Social Psychology Bulletin,* 1-17. [https://doi.org/10.1177/0146167221104231](https://doi.org/10.1177/01461672211042315)Zeelenberg, M., Van den Bos, K., Van Dijk, E., & Pieters, R. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology, 82*(3), 314. https://doi.org/10.1037/0022-3514.82.3.314

Zhang, J. H., Walsh, C., & Bonnefon, J. F. (2005). Between-subject or within-subject measures of regret: Dilemma and solution. *Journal of Experimental Social Psychology, 41*(5), 559-566. <https://doi.org/10.1016/j.jesp.2004.10.004>

Zikmund‐Fisher, B. J., Sarr, B., Fagerlin, A., & Ubel, P. A. (2006). A matter of perspective: choosing for others differs from choosing for yourself in making treatment decisions. *Journal of General Internal Medicine, 21*(6), 618-622. https://doi.org/10.1111/j.1525-1497.2006.00410.x

1. For short-term regrets that are less than a year ago, we estimated based on provided information. For example, for Gilovich and Medvec (1994) Study 5 past-week regret of participants, the estimated no. of “year” is 0.0096 (3.5 days / 365 days, as it is possible for the decision to be 7 days ago or less than a day ago, we picked the middle number). [↑](#footnote-ref-2)
2. We note that for lifetime regrets, we coded the studies with reference to the age of participants, or estimated age based on relevant information (e.g., college students are around 20 years old). In any case, we justified our estimation in the coding sheet. [↑](#footnote-ref-3)
3. The formula for Hedges’ *g* or Cohen’s *d* mayvary among different studies. In a meta-analysis, it is not practical to ensure the effect size measure of all studies having the same meaning or formula across different types of designs. We can only roughly estimate Hedges’ *g* of different studies based on available information (which may be *M*, *SD*, t-statistics, Cohen’s *d*, Chi-Square statistics, proportions of choices) which may be limited in some studies and different across studies. [↑](#footnote-ref-4)
4. As suggested by the reviewer Emiel Cracco, we plan to conduct sensitivity analyses excluding studies in which we are only able to calculate *dz*, as sometimes there can be substantial differences between *dz* and other types of *d*. We plan to report results before exclusion and after exclusion of studies with only t-statistics and sample sizeinformation. [↑](#footnote-ref-5)