

Impact of past behavior normality:
Meta-analysis of exceptionality effect

*Adrien Fillon

Department of Social Psychology
Aix-Marseille University
Aix-en-Provence, France
adrienfillon@hotmail.fr

*Lucas Kutscher

Department of Work and Social Psychology
Maastricht University
Maastricht, 6200MD, the Netherlands
lucaskutscher@gmail.com

*^Gilad Feldman

First affiliation: Department of Psychology
University of Hong Kong
Hong Kong SAR
Second affiliation: Department of Work and Social Psychology
Maastricht University
Maastricht, 6200MD, The Netherlands
gfeldman@hku.hk

In press at Cognition and Emotion

Accepted for publication on August 18, 2020

* Equal contribution

^ Corresponding author

Word: abstract – 179, manuscript - 8406

Corresponding author

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

Author bios:

Adrien Fillon is a graduate of Nanterre University applied social psychology department. His research focuses on open-science and judgment and decision-making.

Lucas Kutscher is a graduate of the Maastricht University work and social psychology department masters program. For his thesis under the guidance of Gilad Feldman he conducted a pre-registered replication and pre-registered meta-analysis of exceptionality effect.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on open/meta-science and judgment decision-making.

Declaration of Conflict of Interest:

The author(s) declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Financial disclosure/funding:

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

Acknowledgments or Author Notes: [None]**Authorship declaration**

Lucas worked under the supervision of Gilad at Maastricht University for conducting the pre-registered meta-analysis as part of Lucas's masters thesis. Lucas wrote the pre-registration, with verification and registration by Gilad. Lucas conducted the search of the literature. Gilad developed the coding scheme. Lucas and Gilad pre-tested the coding sheet. Lucas and Gilad coded the articles. Gilad wrote the RMarkdown code and analyses. Lucas summarized the methods and results and wrote an initial draft. Adrien expanded on the search and coding, reverified and reanalyzed the data, and made major revisions to address feedback received. Gilad and Adrien finalized the manuscript for submission.

Abstract

Exceptionality effect is the phenomenon that people associate stronger negative affect with a negative outcome when it is a result of an exception (abnormal behavior) compared to when it is a result of routine (normal behavior). In this pre-registered meta-analysis, we examined exceptionality effect in 48 studies ($N = 4212$). An analysis of 35 experimental studies ($n = 3332$) showed medium to strong effect ($g = 0.60$, 95% confidence intervals (CI) $[0.41, 0.79]$) for past behavior across several measures (regret/affect: $g = 0.66$, counterfactual thought: $g = 0.39$, self-blame: $g = 0.44$, victim compensation: $g = 0.39$, offender punishment: $g = 0.51$). An analysis of 13 one-sample studies presenting a comparison of exceptional and routine behaviors simultaneously ($n = 1217$) revealed a very strong exceptionality effect (converted $g = 1.98$, CI $[1.57, 2.38]$). We tested several theoretical moderators: norm strength, event controllability, outcome rarity, action versus inaction, and status quo. We found that exceptionality effect was stronger when the routine was aligned with the status quo option and with action rather than for inaction. All materials are available on: <https://osf.io/542c7/>

Keywords: norm theory; normality; regret; past behavior; exception routine; meta-analysis; exceptionality effect

Impact of past behavior normality:

Meta-analysis of exceptionality effect

Kahneman and Miller's (1986) norm theory offered an organizing framework for the role of norms in human cognition and affect and theorized normality as an important factor in feelings of regret. They introduced the concept of normality, broadly referring to the extent to which an event or behavior is perceived as normal (Miller et al., 1990). Norm theory argued that people process normal and abnormal events differently. Normal events are cognitively more salient and readily available than abnormal events, making it easier to simulate alternatives realities (counterfactuals, thoughts of "what might have been") to abnormal events and eliciting stronger affective reactions.

In the present investigation, we report a pre-registered meta-analysis of the impact of past behavior intrapersonal normality on regret. Following arguments by norm theory, evaluations of negative outcomes are affected by a comparison of current behavior to past behaviors: the more exceptional the behavior compared to established past routine the stronger the feelings of regret. We aim to clarify the phenomenon across most common study designs and affect related measures and to examine factors that moderate the effect. We begin with an introduction to the concepts of normality and emotion, then proceed to hypotheses on main effects and potential moderators.

Past behavior normality

Reviews of the literature on norm theory and the concept of normality have so far resulted in the categorization of normality into three broad categories (Feldman & Albarracín, 2017; Koonce et al., 2015): General social norms normality, expectations/situation normality, and past behavior normality. The evidence so far is in support of the three types of normality as distinct categories that are not necessarily related (Feldman, 2020). Much of the research on norm theory and the impact of normality has been conducted on past behavior norms, that is - intrapersonal norms of a focal person (see for example: Gavanski & Wells, 1989; Kahneman & Miller, 1986; Macrae, 1992; Miller & McFarland, 1986), and was therefore chosen to be the focus of this investigation.

We define past behavior normality as the extent to which a person's behavior corresponds with his/her former behavior. Literature refers to past behavior normality (Feldman, 2019; Feldman & Albarracín, 2017; Koonce et al., 2015) by using a variety of different terms such as normal and abnormal behavior (Macrae, 1992; Miller & McFarland, 1986), routine and exception (Davis et al., 1995; Gavanski & Wells, 1989; Lundberg & Frost, 1992; Macrae, 1992; Miller & McFarland, 1986; Turley et al., 1995), intrapersonal normality (Catellani et al., 2004; Hur et al., 2009; McCloy & Byrne, 2000; Roese, 1997) and behavioral consistency (McElroy & Dowd, 2007; Seta et al., 2001, 2008). An overview of terms used for past behavior normality including examples is provided in Table 1.

Past-behavior normality is a core factor in the influential norm theory (Kahneman & Tversky, 1986) and theory of planned behavior (Ajzen, 1991) and has been shown to impact individuals' decisions, affect, and behavior. For example, past-behavior exceptionality impacts both moral emotions and cognition (e.g., regret, blame, compensation), as well as consequences

such as punishment and reward behaviors (Genschow, Rigoni, & Brass, 2017) with implications for real-life legal judgments and behaviors (e.g. Lagnado & Gerstenberg, 2016; Rogers et al., 2010; Wilkinson-Ryan and Baron, 2009). Exceptionality effect has had impact on many other domains, such as folk psychology and morality. For example, Fillon et al. (2020) recently demonstrated links between exceptionality effect and attributions of free will, regret, blame, and compensation when evaluating wrongdoings. The emotion of regret and associated cognitive factor of counterfactual thought are also widely explored factors in the context of past behavior, with important real-life implications for decision-making (Zeelenberg & Pieters, 2007). By conducting meta-analysis of the impact of exceptionality we aimed to examine existing findings to summarize what we already know, define what is yet unclear, identify gaps, and outline future directions regarding what remains unknown.

Routine-exception asymmetry: Past behavior normality and affect

People tend to associate stronger negative feelings with negative outcomes when those are a result of an exceptional compared to routine behavior. Kahneman and Miller (1986) provided the first demonstration of this effect using a thought experiment, referred to as the hitchhiker-scenario:

Mr. Jones almost never takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed.

Mr. Smith frequently takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed.

Who do you expect to experience greater regret over the episode?

Mr. Smith acts as he normally does, whereas Mr. Jones makes an exception to his behavior and deviates from his own established norm. The majority of participants (88%)

answered that Mr. exceptionality Jones is more likely to feel upset about the incident compared to Mr. routine Smith.

This routine-exception asymmetry has been referred by many names, such as emotional amplification (Kahneman & Miller, 1986), normality bias (Koehler & Prentice, 2003), exceptional-routine effect (Catellani et al., 2004; Catellani & Milesi, 2001), exceptionality effect (Byrne, 2016; Dixon & Byrne, 2011; McEleney & Byrne, 2006), and status quo bias (Inman & Zeelenberg, 2002). In this study, we chose to refer to this phenomenon using the term exceptionality effect. We provide an overview of terms for the exceptionality effect including examples in Table 2.

Testing Exceptionality effect: Study designs

The relationship between past behavior and regret has been examined by employing two main study designs. In most early studies participants were presented with two types of behavior simultaneously - exception versus routine - and were asked to compare likely feelings of the two protagonists, as in the hitchhiker-scenario above. Other studies used a between-subject experimental design in which participants were randomly assigned to one of several conditions and were presented with only one behavior type - either exception or routine. Since meta-analyses of several design types are difficult to interpret (Lipsey & Wilson, 2001), we planned separate analyses for each type of study design.

Our main focus was on experimental studies of the exceptionality effect, although we also summarized one-sample comparison studies. We note that one-sample experiments tend to create an artificial contrast between exceptional and routine behavior and are therefore likely to result in inflated effects. As an example of a possible issue with inflated effects, N'Gbala and Branscombe (1997) contrasted different designs in action-inaction regret asymmetries (also

referred to as the *action-effect*; Kahneman & Tversky, 1982). They replicated the action-effect successfully in one-sample studies - when the fate of two persons was presented together, finding stronger regret for action than for inaction. However, the effect was much weaker when using between-subject experimental designs.

Dependent variables.

Studies on exceptionality effect used a diverse set of measures to capture routine-exception asymmetries. We categorized all identified measures into the following: affect, counterfactual thought, self-blame, victim compensation, and offender punishment. Our categorization mainly relies on the original questions that authors used to assess the emotional responses of participants, regardless of the theoretical background of a particular measure. A summary of all included measures and their categorization is provided in Table 3.

Negative affect and regret.

Most studies measured perceived negative affect, like regret, typically asking participants to judge the emotional reaction of the protagonist in a described scenario. They, for example, asked about emotional states, as in the hitchhiker-scenario – “Who do you expect to experience greater regret over the episode?” or the car accident-scenario “Who is more upset over the accident?” (Kahneman & Miller, 1986, p. 145)

Counterfactual thought.

Counterfactual thought was measured in several ways. Some studies used a scale asking about agreement with statements regarding what “should have been” or “would have been” (Roese & Olson, 1996). Other studies measured the number of produced counterfactual alternatives, such as by presenting participants with a scenario with questions regarding events that if had not occurred would undo the negative outcome. Common instructions were “[...] list

six things that could have been different to have changed the outcome” (Gavanski & Wells, 1989, p. 319) or “who spends more time thinking about how things could have been different?” (Briazu et al., 2017, p. 70).

Self-blame.

Some studies referred to self-blame, which is one of the main factors associated with feelings of regret (Connolly & Zeelenberg, 2002). For example, subjects were asked to rate the extent to which “they thought that the [...] victim felt she was to blame” and “they thought that the [...] victim felt responsible” (p. 293).

Victim compensation and Offender punishment.

Some articles used victim compensation and offender punishment as dependent variables in past-behavior exceptionality. In the robbery-scenario (Miller & McFarland, 1986), for example, a man loses the use of his right arm in a robbery taking place in a grocery store, and participants were then asked: “how much money should Mr. Paul receive in compensation for his loss?” (p. 515).

Other articles assessed outcome affect by asking participants to recommend a punishment that an offender should receive for a crime. Turley et al. (1995) for example asked participants to “[...] recommend a prison sentence for the rapist, using a 21-point scale ranging from zero to 20 years” (p. 289).

We note that while compensation and punishment measures are sometimes described as converging, they may prove not to be. We therefore sought to examine whether they were indeed aligned.

Summary of exceptionality effect and measures

We planned a separate analysis for each dependent variable and to also include a moderator analysis comparing the different categories. In line with prior research, we expected that people would feel more negative about an unfortunate outcome that is a result of exceptional behavior in comparison to routine behavior. We expected this exceptionality effect to occur across all types of affect related measures, leading to the following hypotheses:

Hypothesis 1a: People associate stronger negative affect (e.g., regret) with negative outcomes that are a result of exceptional behavior in comparison to routine behavior.

Hypothesis 1b: People engage in more counterfactual thinking following negative outcomes that are a result of exceptional behavior in comparison to routine behavior.

Hypothesis 1c: People associate stronger self-blame with negative outcomes that are a result of exceptional behavior in comparison to usual behavior.

Hypothesis 1d: People assign higher compensation to someone who became the victim of a crime as a result of exceptional behavior in comparison to routine behavior.

Hypothesis 1e: People suggest higher offender punishment when the victim of the crime has shown exceptional behavior in comparison to routine behavior.

We also planned moderator analysis on the type of dependent variable to explore whether the exceptionality effect differs depending on the type of the dependent variable used. We did not have a specific hypothesis regarding the impact of the used measure on the effect size of the main effect.

Theoretical moderators

We expected heterogeneity in the sample to be relatively high and, therefore, planned to conduct moderator analyses to examine moderating factors. Based on the literature, three

moderators were suggested with specific hypotheses, and two other moderators were defined as exploratory. A full description of the hypotheses was included in the pre-registration and is detailed in the supplementary.

Confirmatory moderators

Norm strength.

Past behavior norms can vary in strength. For example, intrapersonal norms are considered strong when they performed as a strict routine, such as *always* walking home on the same route. Strict routine is often described using words such as “always” (Guttentag & Ferrell, 2004; Macrae et al., 1993) or “each day” (Turley et al., 1995). A weaker form of past behavior norms is of usual behavior, for example, when someone *regularly* - but not always - goes shopping at the same store. Usual behavior is often described using the words “often” (Hooker et al., 2000) or “regularly” (Macrae, 1992). Following norm theory logic, an exception would be more abnormal for very strict routines in comparison to repeated behavior. Possibly, the stronger the routine or, the stronger the word or the frequency used to describe past behavior norms, the stronger the exceptionality effect is.

Event controllability.

An exception to past behavior can be made for multiple reasons. People may intentionally deviate from their everyday routines to enrich their lives with excitement and make discoveries. People may also deviate from routines due to uncontrollable circumstances, because they were forced to, or they could not adhere to their routine behavior. For example, Miller and McFarland (1986) used a scenario describing a man who cannot visit his usual grocery store because of ongoing renovations and, therefore, needs to go shopping at another store. In this situation, the

exceptional behavior of shopping at another store was beyond their control, due to an external factor.

Controllable exceptions resulting in negative outcomes may produce stronger regret than uncontrollable ones. According to decision justifiability theory (Connolly & Zeelenberg, 2002), people feel greater regret over decisions that are not justifiable. If a deviation from a past behavior was not intended, then it is easier to justify, and the protagonist has less reason for self-blame or regret because there simply was no other choice to be made, compared to a controllable exception. In line with this argumentation, studies reported the tendency to mutate controllable events as more likely compared to uncontrollable ones (Giroto et al., 1991; N'gbala & Branscombe, 1995). Taken together, anticipated affect likely differs depending on how much control someone had over the exceptional behavior, such that controllable events elicit stronger regret.

Rare versus common outcome.

Our main focus in this meta-analysis is about the distinction between exceptional and routine action in past behavior. Yet, exceptionality is not limited to behavior, but also extends to outcomes. For example, if an unusual accident occurs to a person who followed his/her routine and drove on the commonly chosen road then in terms of behavior this is routine behavior, yet in terms of outcomes, this is an unusual/rare outcome. The first studies to assess mental simulations mainly used scenarios with exceptional rare outcomes like a gunshot in a supermarket (Miller & McFarland, 1986), a car crash, or a robbery by a hitchhiker (Kahneman & Miller, 1986). To our knowledge, the first study to compare common to rare outcomes was by Gavanski and Wells (1989) describing a good student and a bad student receiving a low grade after exceptional versus routine events. They found that participants mainly mutated exceptional behavior in the

direction of normality to undo rare outcomes. We therefore expected exceptionality effect to be stronger for rare outcomes compared to common outcomes.

Summary hypotheses for confirmatory moderators

Our moderator hypotheses were the following:

Moderation hypotheses: Exceptionality effect is stronger (a) the more regular the past behavior was performed, (b) when the exception is controlled compared to when the behavior is uncontrolled, and (c) when the outcome of the behavior is rare compared to when it is a common outcome.

Exploratory Moderators

Action-Inaction.

Action-effect (Kahneman & Tversky, 1982) is the phenomenon that people tend to associate stronger regret with negative outcomes resulting from action compared to inaction. Kahneman and Miller (1986) explained the action-effect using the normality paradigm arguing that action is more abnormal and mutable than inaction and, therefore, regretted more. Subsequent studies have shown support for the impact of normality on the action-effect (e.g., Feldman, 2019; Feldman & Albarracín, 2017; McElroy & Dowd, 2007; Zeelenberg et al., 2002), and that when an action is framed as abnormal, the action-effect is weakened to an *inaction-effect* (Feldman, 2019; Zeelenberg et al., 2002) with higher regret for inaction than for action.

We first expected an exceptionality effect regardless of action-inaction. Further, we entertained competing hypotheses regarding whether the effect would be stronger when the exception is aligned with taking action or with inaction.

Status quo.

Status quo bias (Samuelson & Zeckhauser, 1988) is the phenomenon that people tend to favor a previously chosen or set default over other alternatives. For example, Samuelson and Zeckhauser (1988) demonstrated that when considering financial investments people tend to stick with funds they already hold rather than to invest in alternatives, even when the change was more attractive or profitable.

Routine behavior might be confused for status quo, yet there are important differences between routine behavior and status quo (Feldman et al., 2020). In both cases, routine and status quo are used as reference points, and other options evaluated against them. However, status quo is about sticking to a previously chosen or default option. A status quo situation involves a previous decision favoring a particular choice. Routine, however, represents a mental representation of what previously took place and has over time become the perceived norm, independent of the current decision. We therefore sought to examine whether exceptionality effect would be moderated by the status-quo.

Methods

Pre-registration and open-science

We pre-registered our meta-analysis on the Open Science Framework (OSF) and then proceeded to begin search and coding (<https://osf.io/a65zd>; adjusted to the OSF challenge format: <https://osf.io/hzj3d>). Additional transparency information including disclosures, additional analyses, and key decisions in the meta process are available in the supplementary. These together with the coding sheet, data, and R/RMarkdown code were shared on the OSF (<https://osf.io/542c7/>).

Literature search

We systematically searched Google Scholar (for suitability for meta-analyses see Gehanno et al., 2013; Martín-Martín et al., 2018; Walters, 2007) using fifteen combinations of different keywords without restrictions on the publication date. We first searched for the terms *normality bias*, *exceptional-routine effect* and, *exceptionality effect* which have been used as names for the exception-routine asymmetry. We also combined terms that are related to the independent variable (IV) and the dependent variable (DV). IV related terms were *past behavior*, *past behavior norm*, *past behavior normality*, *abnormal behavior*, *intrapersonal*, *consistency*, *routine-violation*, *behavioral standard*, *behavioral standard*, *past action*, *exceptional events*, *routine events*, *routine*, *typical*, *atypical*, *unusual antecedent*, *usual antecedent*, *exceptional antecedent*, *repeated buying*, *repeated action*, *switching*, *self-blame regret*, *repurchase*, *usual*, *unusual*, *counterfactual version*, *control version* and *prior usage*. All IV searches were run using both behavior and behaviour spelling. DV search terms were *regret*, *upset*, *counterfactual*, *perceived regret*, *anticipated regret*. In every combination of search terms, we integrated the terms *norm theory*, *Miller* and *Kahneman* to narrow the results to literature that has been examined in the context of norm theory.

We provide an overview of the search process in Figure 1. All database searches together achieved 2568 hits. We also searched articles by scanning reference sections of found articles and using the “related articles” and “cited by” option in Google Scholar. All identified sources were imported into a citation management software.

After adjusting for duplicates, 920 sources remained. All abstracts, tables, and results sections of empirical sources were scanned in order to assess their relevance. After this step, 60 articles from 110 unique authors remained as potentially includable articles.

In order to minimize possible potential publication bias, we contacted all identified authors in person and requested unpublished manuscripts. We were provided one additional article leading to a total of 61 sources.

We briefly read through all articles to examine whether they met our inclusion criteria. A total of 46 articles were qualified for exclusion, leading to a total sum of 23 identified articles with codable data. Finally, a total of 48 samples were included in this meta-analysis. We provided a list of all included studies in Table 4.

Coding

Inclusion and exclusion criteria.

Our eligibility criteria addressed five factors: study design, usability of provided data, independent variable, dependent variable, and outcome.

First, we only included experimental studies. Correlational studies were excluded from the sample. Second, sources that did not provide the crucial statistical measures needed for effect size calculations were excluded (such as lacking mean, standard deviation, and/or test statistics). Third, we only included studies that indicated a clear contrast between exceptional and routine past behavior. We excluded all studies that examined other norm types, such as social norms or expectation norms and studies that only measure a deviation from a default (status quo bias) which did not represent repeated behavior or routine. Fourth, we implemented criteria relating to the used affect related measures detailed in Table 3. We excluded measures that were not explicitly related to affect/regret, such as measures of luck, victim positivity, victim sympathy, and victim negligence. Finally, the outcome of a scenario needed to be identified as clearly negative, and we excluded scenarios with neutral or positive outcomes.

Extraction and pre-test.

We developed a data extraction sheet, pilot-tested it on six studies in two stages and refined it accordingly after each stage. The coding process for the pre-tests was completed by two coders to ensure a high inter-rater-reliability. We documented and reported all decisions in detail. After testing, one review author extracted all data and provided detailed information about coding decisions. A second author verified the coding. Disagreements were resolved by discussion between the two authors. All coding decisions were documented in the extraction sheet. The third author again checked and verified all coding.

Included studies.

We included a total of 35 experimental studies and 13 one-sample studies with a total of 4549 participants. The final sample consists of 43 published and 5 unpublished studies. All studies were conducted either in United States ($k = 39$), United Kingdom ($k = 5$), Canada ($k = 3$), or France ($k = 1$). Most studies were conducted in a laboratory context with students as subjects. Some of the more recent studies were online studies using participants other than students. An overview of all included studies is provided in Table 4.

Analysis.

We ran our analysis in R. We used the following meta-analysis related packages to conduct our analyses: metafor, psych, compute.es, MBESS, MAd, powerAnalysis, metaforest, esc, metaviz, puniform. Given the range of different types of studies and experimental designs, we expected heterogeneity in the sample to be relatively high. Therefore, a random effects model was used.

We converted all effect sizes to Cohen's d and standardized into Hedge's g to allow for a comparison. The use of Hedge's g was not pre-registered but was implemented based on feedback from reviewers on early pre-prints to gain an unbiased estimate of the effect size taking

into account sample size. Hedge's g is a correction for the bias in Cohen's d in small samples (mostly $n < 20$) addressing issues of overestimation (Lakens, 2013). The interpretation and use of d and g are similar, as they are identical in samples larger than 20. Split conditions due to moderators in the original studies were collapsed to allow for a comparison of the main independent variable. For within-subject design studies, we used the *agg* function (MAAd package) to aggregate the effects. Because none of the studies reported the correlations between the measures, we used the default correction set to 0.5, following the Wampold et al. (1997) example.

Whenever available, we collected standardized effect sizes directly from authors of original papers. Based on the provided information and details, we checked for the accuracy of these analyses. If effect sizes were not reported, we used either descriptive statistics or inferential statistics to compute standardized effect sizes. All conversions and coding decisions were documented, and the original text was included in the coding sheet to allow for reproducibility.

We produced forest plots of the effect size distribution. A meta-analysis examined the overall main-effect of the bias; a meta-regression was conducted to assess the impact of the described moderators.

Statistical heterogeneity was determined using the Tau² test and quantified using I², which represents the percentage of total variation in a set of studies that is due to heterogeneity (Higgins et al., 2003). This yielded a point estimate, confidence interval, and p-value, along with statistics for heterogeneity, assessed using the Q-statistics, and the I² statistic. We detected significant heterogeneity and therefore proceeded to explore potential moderators.

We also performed analyses for the presence of publication bias, including funnel plots and statistical tests for publication bias (publication status as a moderator) and funnel plot

asymmetry tests (Trim-and-fill method, rank correlation test, Egger's unweighted regression symmetry test, etc.).

Moderator analyses.

We tested subgroups and moderators using a comparison of fixed-effects meta-analysis models. We conducted meta-analyses on subsets of the data, in particular, we split the data by study design and dependent variable types. We tested confirmatory hypotheses for norm strength, event controllability, and outcome rarity (detailed in the moderator section of this paper) and conducted exploratory analyses for action-inaction and status quo as exploratory moderators.

Results

Exceptionality bias main effect

We conducted separate random effects meta-analyses per each of the dependent variables and each of the two study designs. All results are summarized in Table 5 for experimental studies and in Table 6 for one-sample comparison studies.

Experimental studies.

In an analysis of all experimental studies on the impact of past behavior normality on regret, we found an effect of $g = 0.60$ [0.41, 0.79]. We found considerable heterogeneity in the observed effect sizes. 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 also conducted separate analyses for each dependent variable type. Studies measuring regret/affect had an effect of $g = 0.66$ [0.40, 0.92]. Experiments of counterfactual thought ($k = 5$) had an effect of $g = 0.39$ [0.08, 0.70], victim compensation an effect of $g = 0.39$ [0.11, 0.66] offender punishment an effect of $g = 0.51$, [0.30, 0.73], and self-blame an effect of $g = 0.44$ [0.18, 0.69]. Heterogeneity of effects across types was moderate to strong.

One-sample comparison studies.

The second analysis only included comparison one-sample studies and found a very strong effect: $g = 1.98$, [1.57, 2.38]. Regret/affect had an effect of $g = 2.04$, [1.77, 2.32], and counterfactual thought an effect of $g = 1.79$, [-0.43, 4.00] (see details and discussion of these confidence intervals in the supplementary). Tests for heterogeneity revealed moderate to considerable variation in effect sizes across studies.

Study design and measures as moderators.

Results for study design as moderator are provided in Table 7. We expected a stronger exceptionality effect for one-sample studies compared to experimental studies. Indeed, the effect was larger for one-sample studies than for experimental studies. Effect sizes in comparison designs were about three times larger than in experimental studies.

We summarized all other moderator findings in Table 8. We first explored whether exceptionality effect varied between the different dependent measures. The analysis revealed a moderation effect of measures on the main effect for experimental studies. The exceptionality effect was strongest for regret/affect but weaker for counterfactual thought, self-blame, victim compensation, and offender punishment. However, we found no support for differences between the measures when conducting pairwise comparisons and the confidence intervals of effects for regret/affect and the other measures overlapped.

Theoretical moderators

Due to a small number of cases and possibly overestimated effects for one-sample studies (discussed later), we examined theoretical moderators only for experimental designs.

Routine strength.

We hypothesized a stronger exceptionality effect for routine strength as a moderator of exceptionality bias. We found that weaker routine strength is associated with a higher effect size than stronger routine, but the effect was not significant.

Controllability.

We hypothesized stronger exceptionality effect for controllability. We observed higher effect for non-controllable actions than controllable, yet we caution against drawing any conclusions from these findings as we only found one effect size for non-controllability.

Outcome rarity.

We hypothesized that rare outcomes lead to stronger exceptionality effects than common outcomes, but found no support for a moderation effect of this variable.

Action-Inaction.

We tested action-inaction as an exploratory moderator. We found that exceptions aligned with inaction resulted in stronger regret than exceptions aligned with action.

Status Quo.

We explored status quo as a moderator without specific hypotheses. Our findings suggest that the exceptionality effect was stronger when routine was aligned with defaults. However, we identified only three experiments that could be coded as non status-quo, all from a single article (Seta & Seta, 2013), and so we caution against drawing any definite conclusions from this analysis.

Publication bias

We tested for the presence of publication bias using several methods, and a summary of publication bias analyses is provided in Table 9. We ran publication bias analyses on collapsed effect sizes by study, with one effect size per study. Point estimates are fairly consistent, and methods that produce confidence intervals show substantial overlap in confidence intervals for each method. The range of estimates goes from 0.55 to 0.60. The only exception was P-uniform with an estimate of .81. However, methods based on p-values such as P-uniform should not be used to estimate effect sizes with high heterogeneity (>.50%) because they systematically tend to overestimate it (van Aert et al., 2016). The Trim-and-Fill method indicates a possible publication bias with the addition of 8 studies missing on the left. The Trim-and-Fill funnel plot is shown in Figure 3. According to Carter et al. (2019), “the popular trim-and-fill adjustment, while efficient, reduces bias and Type I error rates only slightly. To achieve stronger reductions in bias,

adjustments such as PET-PEESE, p-curve/p-uniform, and 3PSM should be considered” (p.16). Although the likelihood ratio test of 3PSM is significant, all the estimators mentioned by Carter et al. (2019) indicate a moderate to strong effect after adjustment, consistent with the estimate found by the random-effect model. Given that we averaged effect sizes per study we caution of overinterpreting the 3PSM and p-curve/p-uniform analyses which typically assume one effect size per study. Overall, the estimators seem to show no indication for a publication bias in this analysis.

A moderator analysis comparing effects for published ($k = 30$, $g = 0.60$, 95% CI [0.38, 0.82]) and non-published studies ($k = 5$, $g = 0.61$, 95% CI [0.30, 0.92]) in experimental studies did not reveal differences in effect sizes between the two groups (*mean difference*: 0.012, 95% CI [-0.37, 0.39]). We could not conduct this analysis for one-sample studies because all studies were from published articles.

In sum, our analyses revealed no clear support for publication bias in this literature and possible corrections did not cancel the effect.

Discussion

We conducted a meta-analysis and found support for the exceptionality effect. Across different study types and regret/affect measures, we found evidence in support of the phenomenon that people associate stronger reactions with negative outcomes when those are a result of an exception compared to routine.

Exceptionality effect: Main effect

The present meta-analysis found exceptionality effect of medium magnitude for experimental designs ($g = 0.60$) and very large magnitude for one-sample designs ($g = 1.98$). The

effect was consistent across all measures that have been used in the included literature: regret/affect, counterfactual thought, self-blame, compensation, and offender punishment.

Methods moderators

Study design.

The effects were almost three times larger for one-sample studies compared to experimental studies. In one-sample scenarios, participants are presented with the two types of behavior simultaneously (exception and routine) and are asked to compare who of the two protagonists might react more strongly. These thought experiments are powerful illustrations of the bias yet likely overestimate the ‘true’ effect. By presenting exception and routine behaviors together, comparison scenarios likely create an artificial contrast and force the participants to evaluate one against the other. In everyday life, people rarely have the opportunity to compare the fate of two different persons or behaviors directly. In experimental between-subject designs, participants are assigned to conditions and are presented only one type of behavior (exception or routine), which is likely to be more meaningful and realistic for participants. In within-subject designs, participants are presented with both conditions, yet only one type of behavior is evaluated at a time. As mentioned earlier, research on the action-effect has already demonstrated similar differences between experimental and one-sample designs (N’Gbala & Branscombe, 1997).

Dependent variable type.

Exceptionality effect was slightly stronger for studies that measured regret/affect compared to studies that used other measures: counterfactual thought, compensation, offender punishment, or self-blame as dependent variables.

If the effect regarding regret/affect was driven by counterfactual thought alone, and if regret is the sole reason for exceptionality bias for the other dependent variables then we would have observed equal or stronger exceptionality bias for counterfactual thought when compared to regret, and smaller bias for the more distal variables. In our meta-analysis, this was not the case. Regret/affect had the strongest effect size, and counterfactual thought had similar effect sizes as the other variables. One possible explanation is that exceptionality does not affect regret in the same way that it affects other measures. For example, in Miller and McFarland's (1986) article, counterfactual thinking was associated with compensation, but the dependent variable measure was not regret, but rather sympathy for the victim. Recent pre-registered replications of the classic experiment failed to find support for a link between exceptionality and compensation, yet found support for a stronger link between exceptionality and attributions of regret (Fillon et al., 2020; Kutscher & Feldman, 2019). This means that for the exact same paradigm, exceptionality bias can be found for regret but with weaker links to other associated measures like compensation. A promising and important direction for future research is to examine the interplay between the different dependent measures we identified in this meta-analysis, to help uncover the causal chain. For example, the relationship between counterfactual thought and regret is complex, and many studies we found only measured one of the two variables. It is important to further investigate whether regret leads to generating more counterfactual thought (i.e., regret is an emotion that influences cognition, Zeelenberg & Pieters, 2007) or whether counterfactual thought helps to regulate regret (i.e., regret learning, Coricelli & Rustichini, 2010). This may help better understand our finding regarding the differences in effect size for regret and counterfactual thought in exceptionality effect. We see much promise in future research that would help unveil the casual chain linking exceptionality effect with emotions (i.e.,

regret, sympathy), cognition (i.e., counterfactual thought) and behaviors (i.e., blame, praise, and compensation).

Theoretical moderators

We tested five moderators: norm strength, controllability, outcome rarity, action-inaction and status quo. We did not find support for norm strength and controllability as moderators, probably due in part to the small number of studies we were able to find ($k = 5$ for strong norm strength, $k = 1$ for non-controllable action). We found a stronger effect for weak routine than for stronger routine, contrary to our expectations, though we cautioned against drawing any definite conclusions from these analyses. If this generalizes, it might be explained by an interpretation of norm theory that takes into account loss aversion and reference points. According to prospect theory, people weigh losses more heavily than gains of the same magnitude (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991). What is perceived as a loss or gain depends on the reference point. People may consider normal cause as the reference and compare abnormal behavior to this state (Baron, 2004; Ritov & Baron, 1994), and a decision for an abnormal course of action can be seen as a choice that can either lead to a better or a worse outcome than any normal choice. When taking into account loss aversion, potential losses outweigh the gains, so that abnormal behavior is seen as the option with smaller prospects. In line with this paradigm, any routine can serve as a reference that an exception is compared against, no matter how regularly that routine is performed.

In addition, it could be that the differences between high routine and absolute routine are indistinguishable to people so that norms are perceived similarly over a certain threshold, and both of our moderator categories qualified over that threshold.

Outcome rarity.

We initially hypothesized that rare outcomes would lead to a stronger exceptionality effect than common outcomes. We found no support for outcome rarity as a moderator of exceptionality effect.

Lack of support for an effect should not be taken as evidence that there is no effect, yet if indeed there are no differences then one possible explanation to note for future investigations is that outcome rarity moderator is confounded with the dependent variable measures. Experiments that used scenarios with a rare outcome (such as rape and robbery, rare and severe, or food poisoning, rare but less severe) mostly used measures of victim compensation or offender punishment to assess how participants evaluate the crime. In contrast, experiments with lesser rare outcomes tended to measure regret and affect. Regret/affect studies showed stronger exceptionality effect than other measures, and so this moderator analysis has likely been affected by this confound.

Action-Inaction.

Our moderator analysis revealed that the exceptionality effect was stronger when exceptions were aligned with inactions compared to when exceptions were aligned with action. These findings may at first seem to challenge the action-effect (Kahneman & Tversky, 1982), which describes a tendency for people to associate stronger regret over negative outcomes with action than with inaction.

However, a strong exceptionality effect for omission does not necessarily mean that inactions evoke stronger feelings of regret than commissions. Our findings only indicate that the difference between routine regret and exception regret is smaller when the exception is aligned with action compared to when inaction is aligned with an exception.

From an intrapersonal perspective, people whose exceptional behavior is action might simply be less prone to bias. When someone's exceptional behavior is action (and their usual behavior is inaction) that might mean that this person is state-oriented or introverted. A study by McElroy and Dowd (2007) demonstrated that state-oriented people feel strong regrets about both actions and inactions, meaning that deviations from their own routines might not as strongly affect their regrets. In contrast, action-oriented people associate strong regret with inaction (exception) but relatively weaker regrets with action (routine), indicating a strong exception-routine asymmetry. Findings by Seta et al. (2008) showed a similar tendency comparing introverted and extroverted people. Introverts felt strong regret over both action (exception) and inaction (routine) yet extroverts felt relatively lower regret over their commissions. In other words, findings suggest that the more state-oriented or introverted someone is, the less prone they are to exceptionality effect.

Taking a context/social perspective, action-inaction asymmetry may also result from deviations from other reference points - be those social norms or role and situational expectations (Feldman, 2020). A recent review by Feldman, Kutscher, and Yay (2020) linked between action-inaction effects using norm theory, and our findings regarding exceptionality effect are in line and contribute to their integrated model. They outlined several moderators of the action-inaction asymmetry, covering findings which demonstrated that action/inaction asymmetry depends on context and the effect can be weakened and even reversed. For example, nostalgia seems to reverse the effect with stronger regret for inaction than for actions when contemplating the distant past (Gilovich, Medvec, & Kahneman, 1998).

Status Quo.

We found exceptionality effect reversed when a routine was not aligned with the status quo (Seta & Seta, 2013), though we caution that only three studies were coded as non-status-quo, all from a single article.

In most cases, routine represents the default option. Consider, for example, a theater actress who developed the habit to smoke a cigarette before every performance. When she once forgets to have her regular smoke and the show turns out to be a flop, she might then become upset and attribute the failure to the norm violation. Her routine behavior is aligned with her behavioral default and the exceptionality effect would appear as predicted.

However, when someone sets him/herself the goal to change a particular habit, the effect might reverse with stronger regret associated with routine behavior compared to the new but still unusual behavior. Imagine the same actress five years later. She noticed problems with her voice and her doctor, therefore, recommended her to abstain from cigarettes. She set herself the goal to change her default and to refrain from smoking. One night she again feels the need to have her usual cigarette but the performance then becomes a fiasco. She may then attribute her bad performance to her going back to her old habit.

Exceptional behavior might trigger strong regret when the routine represents a default, but the routine might evoke strong regret when someone has set him/herself a new default.

We caution that our findings for this moderator are suggestive given the small number of studies and that they are all associated with a single article and lab (Seta & Seta, 2013). Therefore, the literature would benefit from having more replications and studies examining this research direction on the interaction of status quo and exceptionality bias. Much more research is needed before we can draw any conclusions regarding the impact of status quo.

Implications

Our meta-analysis showed support for the exceptionality effect. This cognitive bias can have a substantial impact on judgments across multiple contexts. These implications might be minor in daily life situations but more severe when it comes to critical decisions in law, business, or politics that have a long-lasting impact on someone.

Studies have shown that people suggested higher victim compensations and more severe prison sentences when circumstances for the victim were exceptional compared to circumstances that were usual (Macrae, 1992; Macrae & Milne, 1992; Macrae et al., 1993; Miller & McFarland, 1986; Turley et al., 1995). The exceptionality of the situation affected the level of harm that was associated with a crime. Ideally, a compensation payment or a punishment should be the product of an objective assessment of the severity of a caused harm and, therefore, not interact with anticipated emotions.

Nevertheless, it should be noted that bias was slightly, but not significantly, stronger for measures of regret and weaker for measures of compensation and punishment. When asked for a judicial decision, people might have been motivated to assess the caused harm more objectively, anticipating the role of a judge or jury in a process. However, the exceptionality effect persisted. Therefore, we emphasize that every decision-maker should be aware of the bias to avoid severe mistakes in key decisions.

Limitations

Sample size.

One limitation of this meta-analysis is the small number of included studies. Since the beginning of norm theory (Kahneman & Miller, 1986), there has been an impressive number of papers that followed and built on this work. However, the number of studies that clearly contrasted routine behavior and exception remains relatively small, with most of those studies

using small samples. Only seven included studies used more than 100 participants. This leads to relatively low statistical power and, therefore, allows only limited conclusions about the durability of the effect (Lakens & Evers, 2014; Maxwell, Lau, & Howard, 2015). All findings should, therefore, be interpreted with caution. This can be taken as an opportunity for future research to employ large samples in tackling this research direction.

Cultural homogeneity.

Our analysis consists of culturally homogeneous samples and, therefore, allows no strong conclusions regarding whether the effect persists across cultures. All studies have been conducted in the United States, the United Kingdom, Canada, or France – countries that broadly share common cultural values. Yet, cultural differences may be a potential moderator of the effect. Hur et al. (2009) demonstrated that cultural orientation could impact the relationship between norms and regret. In their study, Koreans (but not Americans) felt stronger regret over violations of intrapersonal norms compared to deviations from interpersonal norms.

Hofstede (1984) introduced a typology of five dimensions that outline differences between cultures. One of these dimensions is uncertainty avoidance (UA), describing the extent to which members of a culture feel the need to control uncertainties and future events. Making an exception and leaving the familiar path of a routine contains a momentum of uncertainty. Someone who has a high need to avoid uncertainty might be even more regretful over unfortunate happenings that are a result of exceptions, compared to someone who can handle uncertainty well. Consequently, the exceptionality effect may be stronger in countries with high UA. In these countries, the deviation from past behavior would at the same time represent a violation of a social norm, which has also shown to amplify regret (Feldman & Albarracín, 2017; Koehler & Prentice, 2003).

Coding one-sample studies.

The meta-analysis is also limited by one of our coding procedures. Some one-sample studies included a “not sure/same” option that a participant could choose when they were not sure which protagonist might feel worse. We excluded these participants from the sample to create a clear contrast between exception and routine options. This coding procedure is problematic for two reasons. First, it reduces the number of participants of a study and, therefore, inflates effect sizes of comparisons even further. Second, when a single sample participant responded to several comparison scenarios, we had to exclude a different number of participants for each of the within-person responses. This asymmetry across measures created problems when collapsing subgroups in the analysis to assess the overall effect for the entire sample. We have no simple suggestions for overcoming this issue in future research, yet this should be noted and addressed in future metas.

Approximations of effects.

Meta-analyses are a common method to obtain a relatively reliable estimate of an effect, yet their value highly depends on the quality of their input. In particular, when working with earlier papers we were faced with several challenges. Some studies, for example, only provided the total number of study participants and did not communicate their distribution across conditions. Other papers reported results only for particular groups. For this reason, we were often required to work with assumptions and approximations to estimate an effect what finally, what consequently leads to less accurate results. This drawback may apply to almost every meta-analysis in empirical sciences, yet it cannot be overemphasized how severely meta-analyses suffer from incomplete reporting and low transparency.

Directions for future research

Experimental studies.

Our analysis found almost three times larger effect sizes for one-sample studies compared to experimental studies. We argued that comparison scenarios overestimate the ‘true’ effect sizes by creating an unrealistic contrast. One-sample studies illustrate a bias and can be taken as initial evidence that bias might exist but may offer only limited insights regarding the magnitude of the effect. Future work can directly contrast the two designs to directly examine these effects. Till then, we consider experimental studies to be more realistic and therefore recommend future studies to mainly focus on experimental designs.

Call for more pre-registered replications and/or registered replications reports.

One considerable weakness of the present analysis is the relatively small number of studies that have investigated past behavior exceptionality effect. More studies are needed to conclude how robust the effect is. Collaborative replication attempts might help to form a more generalizable statement about the durability of the effect (Maxwell et al., 2015), in particular when they are performed across different contexts and cultural environments. A larger pool of studies might also further clarify the impact of moderating factors. Only two of the studies involved were pre-registered (Fillon et al., 2020; Kutscher & Feldman, 2019). Their results show effects in the range of the effects of the meta-analysis and with narrow confidence interval ($g = .30$ to $.84$). Therefore, in the context of exceptionality, and more broadly for norm theory, pre-registered well-powered replications suggest medium to strong effects (e.g., Feldman et al., 2020: $d = .51$ to $.85$).

Explore moderating factors.

In addition to replications, future work may also aim to further clarify the underlying mechanisms of the exceptionality effect. Our analysis only had limited explanatory power

regarding moderation effects due to the relatively small number of cases, as for example we only found one study for the non-controllability condition. We also found too few experiments to allow for meaningfully differentiating between exceptionality and severity.

Also, we aggregated very weak and medium strength routine into a “weak routine” condition to compare with strong routine, yet found only five studies matching this criterion. To be sure that participants understand the differences between normality and exceptionality, researchers have emphasized the normal condition, using terms such as “always”, “never”, “each day”, “every day/evening”, even sometimes saying that the character is “a creature of habits”. There have not been many studies examining medium to weak strength, to examine norm strength. Normality strength seems like a promising direction for future research on norm theory.

Finally, norm theory has helped explain many biases in the literature and predict decision-making behaviors in various domains. Yet the “normal” in normality was left ambiguous: what is the normality based on? In our analysis, we focused on normality based on past-behavior. Other types of normality include normality based on norms, or on expectation/role. The type of normality has begun to be put in place recently (Feldman & Albarracín, 2017).

The literature would benefit from a systematic elaboration of norm theory examining the many ambiguities regarding normality and factors that might affect normality. We took the first step to examine several factors in our meta-analysis with respect to past-behavior normality and aimed to clarify the literature and examine related moderators. We hope to see more work done in the area to create a typology of normality categories and a theoretical model that would help summarize the literature and point for future research directions. Normality is an important

concept in decision making, and much more work is needed to better understand this concept and its implications.

(*references included in the meta-analysis are marked with an asterisk at the end)

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
- Ball, R. A. (2004). An investigation of biased depictions of normality in counterfactual scenario studies. University of Victoria. <http://dspace.library.uvic.ca:8080/handle/1828/546> *
- Baron, J. (2004). *Normative Models of Judgment and Decision Making*. In Blackwell Handbook of Judgment and Decision Making (pp. 19–36).
<https://doi.org/10.1002/9780470752937.ch2>
- Briazu, R. A., Walsh, C. R., Deeprose, C., & Ganis, G. (2017). Undoing the past in order to lie in the present: Counterfactual thinking and deceptive communication. *Cognition*, 161, 66–73. *
- Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, 67, 135–157.
- 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>
- Catellani, P., Alberici, A. I., & Milesi, P. (2004). Counterfactual thinking and stereotypes: The nonconformity effect. *European Journal of Social Psychology*, 34, 421–436.
- Catellani, P., & Milesi, P. (2001). Counterfactuals and roles: Mock victims' and perpetrators' accounts of judicial cases. *European Journal of Social Psychology*, 31, 247–264.
- Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science*, 11, 212–216.

- 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.
- Dixon, J. E., & Byrne, R. M. (2011). “If only” counterfactual thoughts about exceptional actions. *Memory & Cognition*, 39, 1317–1331.
- Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: evidence for two conceptual systems. *Journal of Personality and Social Psychology*, 62, 328–339. *
- Feldman, G. (2020). What is normal? Dimensions of action-inaction normality and their impact on regret in the action-effect. *Cognition and Emotion*, 1–15.
<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.
- Feldman, G., Kutscher, L., & Yay, T. (2018). What is action, what is inaction? A review of action-inaction biases and recommendations for term use and typology. Unpublished. <https://doi.org/10.13140/RG.2.2.27961.83046>
- Feldman, G., Kutscher, L., & Yay, T. (2020). Omission and commission in judgment and decision making: Understanding and linking action-inaction effects using the concept of normality. *Social and Personality Psychology Compass*. Conditionally accepted. Preprint retrieved from

https://www.researchgate.net/publication/328137078_Omission_and_commission_in_judgment_and_decision_making_Understanding_and_linking_action-inaction_effects_using_the_concept_of_normality

- Fillon, A. A., Lantian, A., Feldman, G., & N'Gbala, A. (2020). Exceptionality Effect in Agency: Exceptional Choices Attributed Higher Free Will Than Routine. *Center for Open Science*.
<https://doi.org/10.31234/osf.io/sep8f>*
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314–325. *
- Genschow, O., Rigoni, D., & Brass, M. (2017). Belief in free will affects causal attributions when judging others' behavior. *Proceedings of the National Academy of Sciences*, 114(38), 10071–10076. <https://doi.org/10.1073/pnas.1701916114>
- Gehanno, J.-F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, 13, 1–5.
- Gilovich, T., Medvec, V. H., & Kahneman, D. (1998). Varieties of regret: A debate and partial resolution. *Psychological Review*, 105(3), 602.
- Giroto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78, 111–133.
- Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: age differences in judgments of regret and relief. *Developmental Psychology*, 40, 764–775. *
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, 327-557.
- Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of*

- Management*, 1, 81–99.
- Hooker, C., Roese, N. J., & Park, S. (2000). Impoverished counterfactual thinking is associated with schizophrenia. *Psychiatry*, 63, 326–335. *
- Hur, T., Roese, N. J., & Namkoong, J.-E. (2009). Regrets in the East and West: Role of intrapersonal versus interpersonal norms. *Asian Journal of Social Psychology*, 12, 151–156.
- Inman, J. J., & Zeelenberg, M. (2002). Regret in repeat purchase versus switching decisions: The attenuating role of decision justifiability. *Journal of Consumer Research*, 29, 116–128. *
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263. doi:10.2307/1914185
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136–153. doi:10.1037/0033-295x.93.2.136 *
- Kahneman, D., & Tversky, A. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, 59(4), 251–278.
- Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246, 160–173.
- Koehler, J. J., & Prentice, R. A. (2003). A normality bias in legal decision making. *Cornell Law Review*, 88, 583–645.
- Koonce, L., Miller, J., & Winchel, J. (2015). The effects of norms on investor reactions to derivative use. *Contemporary Accounting Research*, 32, 1529–1554. *
- Kutscher, L., & Feldman, G. (2019). The impact of past behavior normality on regret: Replication and extension of three experiments of the exceptionality effect. *Cognition and Emotion*. 10.1080/02699931.2018.1504747 *

- Lagnado, D. A., & Gerstenberg, T. (2016). Causation in Legal and Moral Reasoning. In M. R. Waldmann (Ed.), *Oxford Handbooks Online*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199399550.013.30>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4.
<https://doi.org/10.3389/fpsyg.2013.00863>
- Lakens, D., & Evers, E. R. (2014). Sailing from the seas of chaos into the corridor of stability practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, 9, 278–292.
- Landman, J. (1987). Regret and elation following action and inaction: Affective responses to positive versus negative outcomes. *Personality and Social Psychology Bulletin*, 13, 524–536. *
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Sage publications.
<http://rogeriofvieira.com/wp-content/uploads/2016/05/Wilson.pdf>
- Lundberg, C. G., & Frost, D. E. (1992). Counterfactuals in financial decision making. *Acta Psychologica*, 79, 227–244.
- Macrae, C. N. (1992). A tale of two curries: Counterfactual thinking and accident-related judgments. *Personality and Social Psychology Bulletin*, 18, 84–87. *
- Macrae, C. N., & Milne, A. B. (1992). A curry for your thoughts: Empathic effects on counterfactual thinking. *Personality and Social Psychology Bulletin*, 18, 625–630. *
- Macrae, C. N., Milne, A. B., & Griffiths, R. J. (1993). Counterfactual thinking and the perception of criminal behavior. *British Journal of Psychology*, 84, 221–226. *
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google

- Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. <https://doi.org/10.31235/osf.io/42nkm>
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, 70, 487–498.
- McCloy, R., & Byrne, R. M. (2000). Counterfactual thinking about controllable events. *Memory & Cognition*, 28, 1071–1078.
- McEleney, A., & Byrne, R. M. (2006). Spontaneous counterfactual thoughts and causal explanations. *Thinking & Reasoning*, 12, 235–255.
- McElroy, T., & Dowd, K. (2007). Action orientation, consistency and feelings of regret. *Judgment and Decision Making*, 2, 333–341. *
- Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, 12, 513–519. *
- Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: Thinking about what might have been. *Advances in Experimental Social Psychology*, 23, 305–331.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6, e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- N’gbala, A., & Branscombe, N. R. (1995). Mental simulation and causal attribution: When simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, 31, 139–162.
- 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, 324–343.
- Ritov, I., & Baron, J. (1994). Judgements of compensation for misfortune: The role of expectation. *European Journal of Social Psychology*, 24(5), 525–539.
<https://doi.org/10.1002/ejsp.2420240502>
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133–148.
- Roese, N. J., & Olson, J. M. (1996). Counterfactuals, causal attributions, and the hindsight bias: A conceptual integration. *Journal of Experimental Social Psychology*, 32, 197–227. *
- Rogers, P., Davies, M., Anderson, I., & Potton, A. (2010). Can Norm Theory Explain the Effects of Victim Age and Level of Physical Maturity on Perceptions of Child Sexual Abuse? *Journal of Interpersonal Violence*, 26(10), 1911–1933.
<https://doi.org/10.1177/0886260510372936>
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1, 7–59.
- Seta, C. E., & Seta, J. J. (2013). Regret in pursuit of change and maintenance goals. *Motivation and Emotion*, 37, 177–184. *
- Seta, C. E., Seta, J. J., McElroy, G. T., & Hatz, J. (2008). Regret: The roles of consistency-fit and counterfactual salience. *Social Cognition*, 26, 700–719. *
- Seta, J. J., McElroy, T., & Seta, C. E. (2001). To do or not to do: Desirability and consistency mediate judgments of regret. *Journal of Personality and Social Psychology*, 80(6), 861–870. <https://doi.org/10.1037/0022-3514.80.6.861>
- Turley, K. J., Sanna, L. J., & Reiter, R. L. (1995). Counterfactual thinking and perceptions of rape. *Basic and Applied Social Psychology*, 17, 285–303. *

- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106, 1039–1061.
- Valshtein, T. J., & Seta, C. E. (2016). Regret for some, regret for all: consistency, outcome severity, and feelings of regret. San Diego. *
- 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(5), 713-729.
- Walters, W. H. (2007). Google Scholar coverage of a multidisciplinary field. *Information Processing & Management*, 43, 1121–1132.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis 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>
- Wilkinson-Ryan, T., & Baron, J. (2009). Moral judgment and moral heuristics in breach of contract. *Journal of Empirical Legal Studies*, 6(2), 405-423.
<https://doi.org/10.1111/j.1740-1461.2009.01148.x>
- Yay, T. (2017) Action-inaction in attributions of harm and blame: a pre-registered replication and meta-analysis. Unpublished document.
https://osf.io/vdb3k/?view_only=91512350e10747b0a4767d92339b30ff
- 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, 314–327.
- Zeelenberg, M., & Pieters, R. (2007). A Theory of Regret Regulation 1.0. *Journal of Consumer Psychology*, 17(1), 3–18. https://doi.org/10.1207/s15327663jcp1701_3

Tables

Table 1

Terms used for past behavior normality

Term	Literature	Example
Normal vs. abnormal	Macrae (1993) Miller & McFarland (1986) Kahneman & Miller (1986) Turley (1995)	“We examine this question in conjunction with a hypothesis of emotional amplification, which states that the affective response to an event is enhanced if its causes are abnormal ” ^a
Exception vs. routine	Kahneman & Miller (1986) Gavanski & Wells (1989) Macrae 1992 Turley (1995)	"[...] the same undesirable outcome is judged to be more upsetting when the action that led to it was exceptional than when it was routine " ^b
Intrapersonal normality	Hur et al. Catellani et al. McCloy & Byrne (2000) Roese (1997)	“[...] intrapersonal norms , which center on a person’s behavior and what that person typically does, [...]” ^c
Behavioral standard	Catellani et al. (2004)	“[...] behaviors deviating from the target's own behavioral standard (intrapersonal norm) are likely to trigger counterfactuals” ^d

Term	Literature	Example
Behavioral consistency	McElroy & Dowd (2007) C. E. Seta et al. (2008) J. J. Seta et al. (2001)	“The consistency-fit conceptualization of regret is that it is a negative emotional state that results from a mismatch between a poor decision and the decision-maker's goals, mood-state or personality characteristics, as well as from the consequences that are associated with a decision that turns out poorly. When there is a mismatch, and thus a poor consistency-fit, decision-makers experience more regret than when there is not a mismatch.” ^e
Past behavior normality	Koonce et al. (2015) Feldman & Albarracín (2017) Feldman (2017)	“Normal can be evaluated using several types of normality [...], most notably – (1) the extent to which a behavior is similar to past behavior [...]” ^f

Note. ^{a, b}Kahneman & Miller (1986, p. 145), ^cHur et al. (2009, p. 152), ^dCatellani et al. (2004, p. 421), ^eSeta et al. (2008, p. 6), ^f(Feldman & Albarracín, 2017, p. 2)

Table 2

Terms used for the exceptionality effect

Name	Literature	Example
Emotional amplification	Kahneman & Miller (1986)	"We examine this question in conjunction with a hypothesis of emotional amplification , which states that the affective response to an event is enhanced if its causes are abnormal." ^a
Normality bias	Prentice & Koehler (2003)	"Because Mr. Y was mugged while taking an unusual route home, this crime will be judged more severely than the mugging of Mr. X, which took place in the context of the usual route. This is an example of what we refer to as the normality bias - the tendency for people to react more strongly to bad outcomes that spring from abnormal circumstances than to otherwise identical outcomes that spring from more ordinary circumstances." ^b
Exceptional-routine effect	Catellani et al. (2004) Catellani & Milesi (2001)	"With regards to the exceptional-routine effect , several studies have shown that in counterfactual simulation elements that are perceived to be exceptional or abnormal are more likely to be mutated than elements that are perceived to be normal." ^c

Name	Literature	Example
Status quo effect	Inman & Zeelenberg (2002)	“[...] we use the term status quo effect in this article to describe consumers’ tendency to feel more regret following a decision to switch than following a decision to repeat a past decision (e.g., to choose the same brand).” ^d
Exceptionality effect	Byrne (2016) Dixon & Byrne (2011) McEleney & Byrne (2006)	“ Exceptionality effect : the tendency to imagine an alternative by changing exceptional events to be normal.” ^e

Note. ^aKahneman & Miller (1986, p. 145), ^bPrentice & Koehler (2003, p. 587), ^cCatellani, Alberici, & Milesi (2004, p. 422), ^dInman & Zeelenberg (2002, p. 116), ^eByrne (2016, p. 713)

Table 3

Definitions and measures of variables in the meta-analysis

Measure group	Included measure	Definition
Affect/Regret (R)	Regret	The extent to which someone feels regret over a negative outcome
	Negative affect	The extent to which someone is upset or foolish about a situation
Counterfactual thought (CF)	Counterfactuals	The number of alternative versions of the past someone imagines facing a given outcome
	Mutation rank	The extent to which someone prioritizes to change a forgone event
	Avoidability	The extent to which someone believes that a situation could have been avoided
Self-blame (SB)	Self-blame	The extent to which someone is blaming him or herself for a negative outcome
	Victim responsibility	The extent to which victim of a crime feels responsible for an outcome
	Victim fault	The extent to which victim of a crime feels to be fault for an outcome
Victim compensation (VC)	Victim compensation	The amount of recommended compensation for a crime victim
Offender punishment (OP)	Prison sentence	The severity of a recommended sentence or fine for a perpetrator
	Fine	The severity of a recommended fine for a perpetrator

Table 4

All Studies included in the meta-analysis

	Article	<i>N</i>	Design	Measures	Published
1	Ball (2004)	67	Within	R, VC, SB, OP, CF	No
2	Ball (2004)	207	Between	OP	No
3	Briazu et al. (2017)	81	One-sample	CF	Yes
4	Epstein et al. (1992)	72	One-sample	R	Yes
5	Epstein et al. (1992)	107	One-sample	R	Yes
6	Fillon et al. (2020)	561	Between	R	No
7	Fillon et al. (2020)	128	Between	R	No
8	Gavanski & Wells (1989)	57	Between	CF	Yes
9	Gavanski & Wells (1989)	57	Between	CF	Yes
10	Guttentag & Ferrell (2004)	4	One-sample	R	Yes
11	Guttentag & Ferrell (2004)	23	One-sample	R	Yes
12	Guttentag & Ferrell (2004)	50	One-sample	R	Yes
13	Guttentag & Ferrell (2004)	71	One-sample	R	Yes
14	Hooker et al. (2000)	41	One-sample	R	Yes
15	Inman & Zeelenberg (2002)	68	Between	R	Yes
16	Inman & Zeelenberg (2002)	68	Between	R	Yes
17	Kahneman & Miller (1986)	92	One-sample	R	Yes
18	Kahneman & Miller (1986)	138	One-sample	R	Yes
19	Koonce et al. (2015)	93	Between	SB	Yes
20	Koonce et al. (2015)	93	Between	SB	Yes
21	Kutscher & Feldman (2019)	347	One-sample	R, VC	Yes
22	Landman (1987)	48	One-sample	R	Yes
23	Macrae & Milne (1992)	30	Within	VC, OP	Yes
24	Macrae (1992)	40	Between	VC, OP	Yes
25	Macrae (1992)	40	Between	VC, OP	Yes
26	Macrae et al. (1993)	50	Between	VC	Yes
27	McElroy & Dowd (2007)	56	Between	R	Yes
28	McElroy & Dowd (2007)	65	Between	R	Yes

29	McElroy & Dowd (2007)	100	Between	R	Yes
30	McElroy & Dowd (2007)	74	Between	R	Yes
31	Miller & McFarland (1986)	163	Between	VC	Yes
32	Roese & Olson (1996)	148	Between	CF	Yes
33	Seta et al. (2001)	40	Between	R	Yes
34	Seta et al. (2001)	39	Between	R	Yes
35	Seta et al. (2001)	38	Between	R	Yes
36	Seta et al. (2008)	73	Between	R	Yes
37	Seta et al. (2008)	73	Between	R	Yes
38	Seta et al. (2008)	60	Between	R	Yes
39	Seta et al. (2013)	59	Between	R	Yes
40	Seta et al. (2013)	60	Between	R	Yes
41	Seta et al. (2013)	67	Between	R	Yes
42	Seta et al. (2013)	67	Between	R	Yes
43	Seta et al. (2013)	45	Between	R	Yes
44	Seta et al. (2013)	45	Between	R	Yes
45	Turley et al. (1995)	62	Between	CF, SB, R, VC, OP	Yes
46	Turley et al. (1995)	80	Between	CF, SB, R, OP	Yes
47	Turley et al. (1995)	66	Between	SB, R	Yes
48	Valshtein & Seta (2016)	167	Between	R	No

Note. Abbreviations for grouped measures: R = Regret, CF = Counterfactual thought, VC= Victim compensation, SB = Self-Blame, OP = Offender punishment

Table 5

Meta-analysis results for the relationship between past behavior normality and regret measures in experimental studies

Measure	<i>k</i>	<i>N</i>	Mean <i>g</i>	95% CI	<i>Q</i>	df	<i>p</i>	<i>I</i> ²
Regret	25	2502	0.66	[0.40, 0.92]	158.89	24	< .001	.91
Counterfactual thought	5	323	0.39	[0.08, 0.70]	8.48	4	.08	.54
Victim compensation	6	687	0.39	[0.11, 0.66]	11.69	5	.04	.60
Offender punishment	8	576	0.51	[0.30, 0.73]	11.34	7	< .001	.42
Self-blame	6	461	0.44	[0.18, 0.69]	11.98	5	< .001	.58
Combined	35	3332	0.60	[0.41, 0.79]	199.33	34	< .001	.88

Note. *k* = number of samples; *N* = total number of individuals in *k*; mean *g* = average Hedge's *g* effect size, CI = lower and upper limits of 95% confidence interval

Table 6

Meta-analysis results for the relationship between past behavior norms and regret measures in one-sample comparison studies

Measure	<i>k</i>	<i>N</i>	Mean <i>g</i>	<i>Q</i>	df	<i>p</i>	<i>I</i> ²
Regret	11	988	2.04	24.52	10	<.01	.53
Counterfactual thought	2	229	1.79	34.17	1	<.001	.97
Combined	13	1217	1.98	94.81	12	<.001	.85

Note. *k* = number of samples; *N* = total number of individuals in *k*; mean *g* = average converted Hedge's *g* effect size, CI = lower and upper limits of 95% confidence interval

Table 7

Results of moderator analysis for study designs

Moderator	<i>k</i>	<i>N</i>	Mean <i>g</i>	95% CI	Difference	<i>p</i>
Design						
One-sample (Comparison)	13	1217	1.98	[1.57, 2.38]		
Experimental	35	3332	0.60	[0.41, 0.79]	-1.38* [-1.83, -0.93]	< .001

Note. *k* = number of samples; *N* = total number of individuals in *k*; mean *g* = average Hedge's *g* effect size, CI = lower and upper limits of 95% confidence interval,

** *p* < .001, two-tailed

Table 8

Moderator analysis for experimental studies

Moderator	<i>k</i>	<i>N</i>	Mean <i>g</i>	95% CI	<i>Difference</i>	<i>p</i>
Measure						
Regret	25	2502	0.66	[0.40, 0.92]		
Counterfactual thought	5	323	0.39	[0.08, 0.70]	-0.30 [-0.93, 0.33]	.35
Self-blame	6	461	0.44	[0.18, 0.69]	-0.40 [-0.65, -0.16]	.001
Victim compensation	6	687	0.39	[0.11, 0.66]	-0.34 [-0.67, -0.02]	.04
Offender punishment	8	576	0.51	[0.30, 0.73]	-0.28 [-0.55, -0.05]	.05
Routine strength						
Medium	29	3476	0.61	[0.39, 0.83]		
Strong	5	297	0.42	[0.08, 0.76]	0.19* [-0.22, -0.59]	.36
Controllability						
Uncontrollable	1	167	1.00	[0.68, 1.32]		
Controllable	16	1609	0.67	[0.50, 0.85]	0.33 [-0.04, 0.69]	.08
Outcome rarity						
Common outcome	13	917	0.77	[0.51, 1.03]		
Rare outcome	14	1899	0.55	[0.41, 0.69]	0.22 [-0.08, 0.52]	.15
Action-inaction						
Inaction	6	443	1.09	[0.87, 1.31]		
Action	21	2505	0.58	[0.43, 0.73]	0.51*** [0.25, 0.78]	<.001
Status Quo						
Status Quo	30	3047	0.73	[0.58, 0.87]		
Non-Status Quo	3	171	-0.76	[-1.06, -0.45]	1.49*** [1.15, 1.82]	<.001
Outcome severity						
High severity	13	1841	0.55	[0.40, 0.70]		
Low severity	23	2052	0.63	[0.34, 0.92]	0.09 [-0.25, 0.41]	.62

Note. *k* = number of samples; *N* = total number of individuals in *k*; mean *g* = average Hedge's *g* effect size, CI = lower and upper limits of 95% confidence interval,

* *p* < .05, two-tailed

***p* < .01, two-tailed

*** *p* < .001, two-tailed

Table 9

Publication biases analyses results

Publication bias analysis method	Results and adjusted models
Three-parameter selection model	Likelihood Ratio Test: 9.22, $p = .002$ Adjusted Model: $g = 0.60$, 95% CI [0.41, 0.79]
PET	$b = 0.59$ [0.21, 0.96], $p < .01$
PEESE	$b = 0.55$ [0.31, 0.80], $p < .001$
Puniform	Adjusted Model: $g = 0.81$, 95% CI [0.67, 0.93], 26 significant
Henmi & Copas (2010)	Adjusted Model: $g = 0.60$, 95% CI [0.41, 0.79]
Trim and fill funnel plot asymmetry	Eight studies missing on the left side.
Rank correlation test	
(Begg & Mazumdar, 1994)	Kendall's tau = 0.14, $p = .24$
Egger's regression test	$z = 1.06$, $p = .288$

Note. Values in parentheses indicate 95% confidence intervals [lower bound, upper bound]. We ran publication bias analyses on collapsed effect sizes by study, leading to one effect size per study.

Figures

Figure 1

Meta-analysis flow diagram (adapted from PRISMA 2009, Moher et al., 2009)

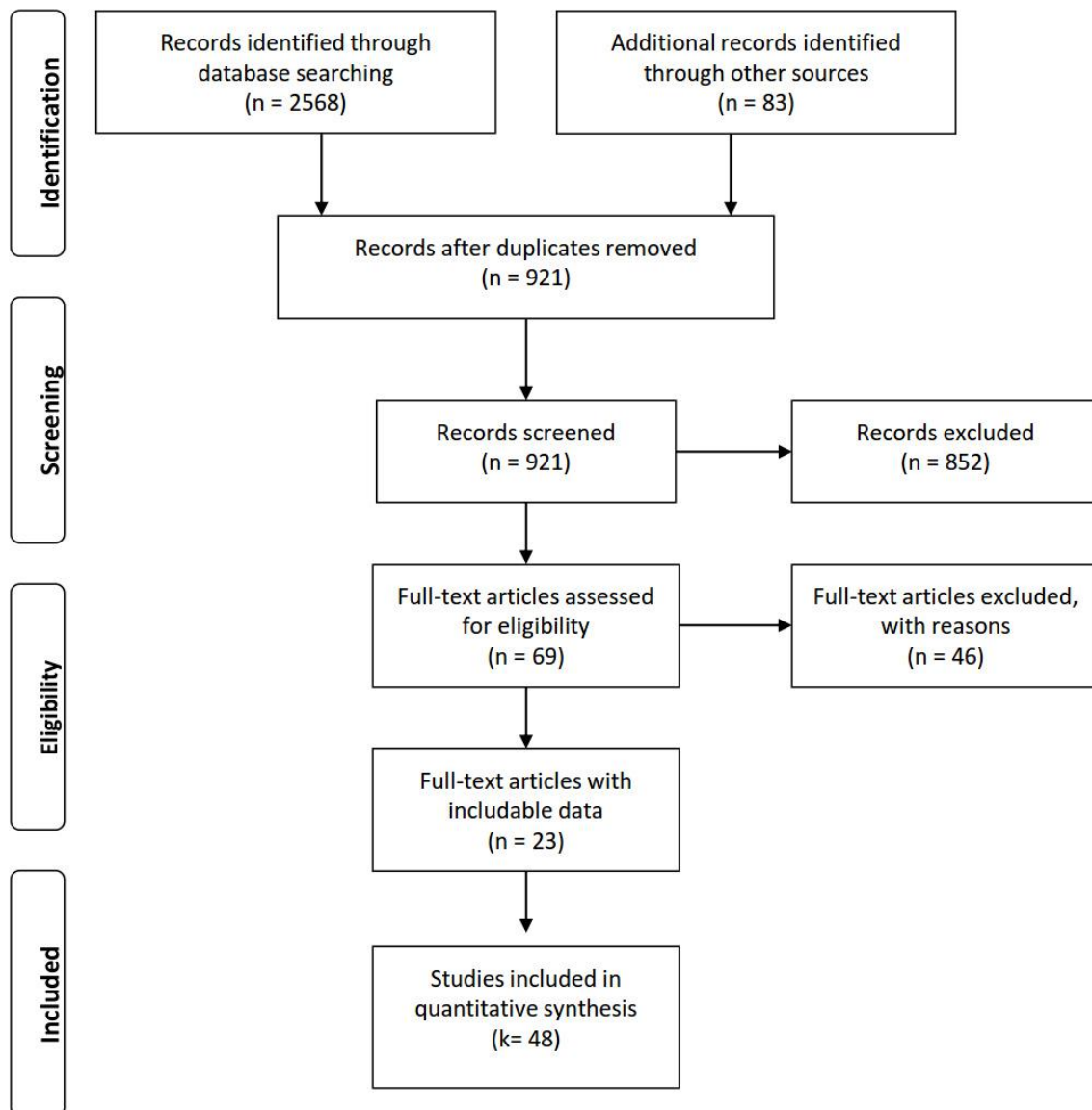


Figure 2

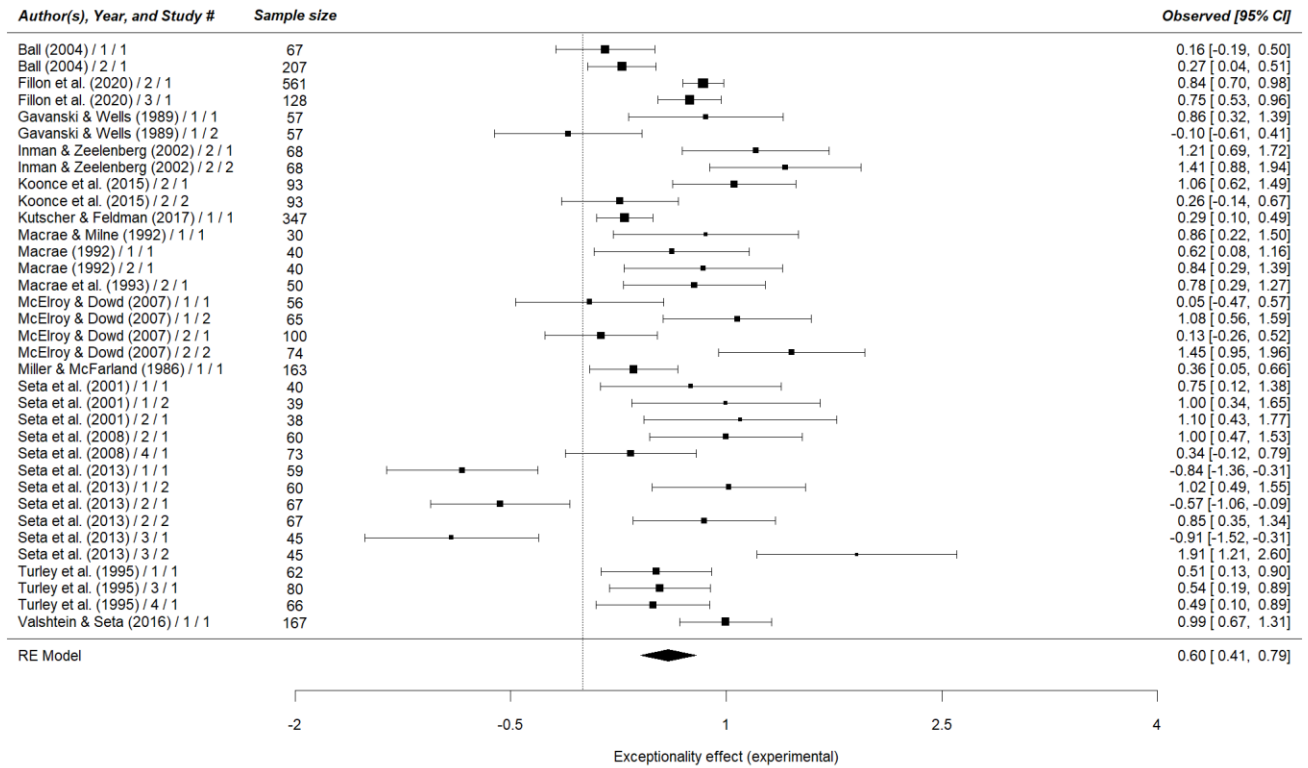
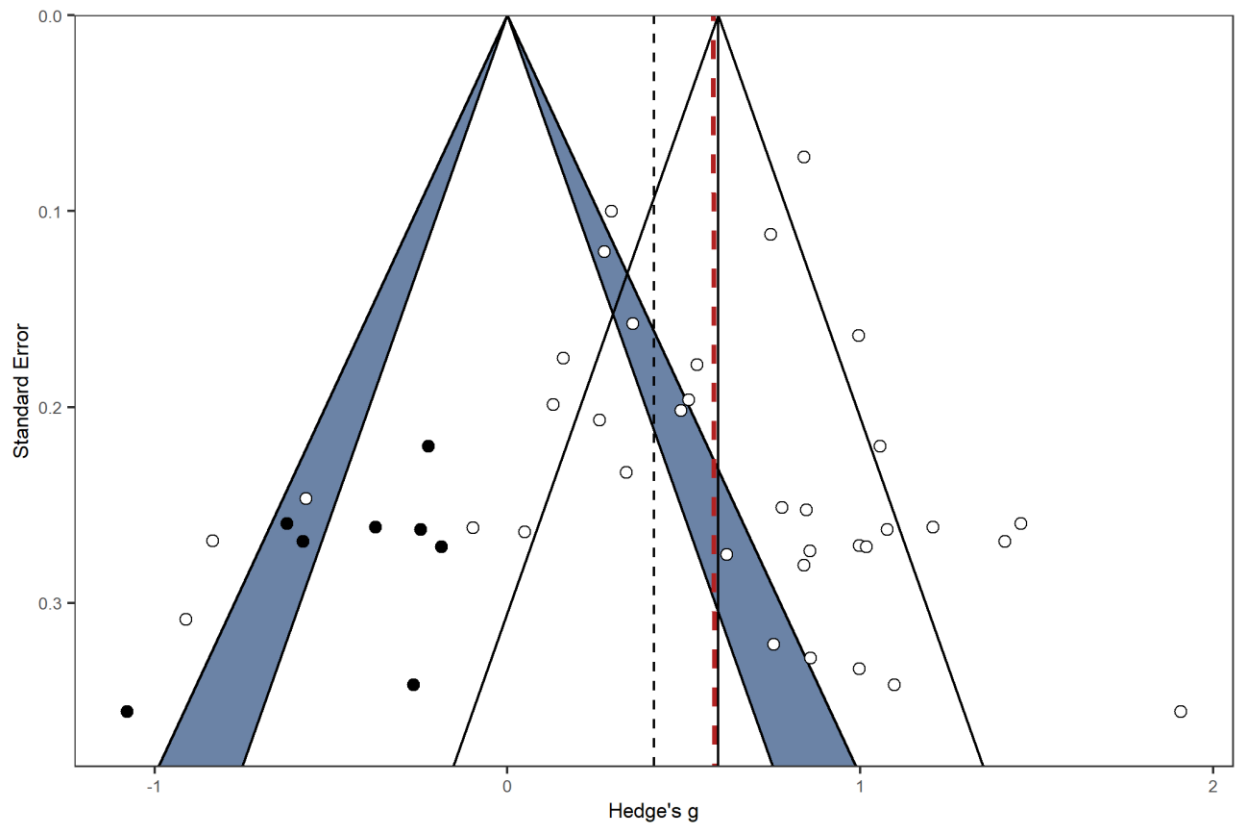
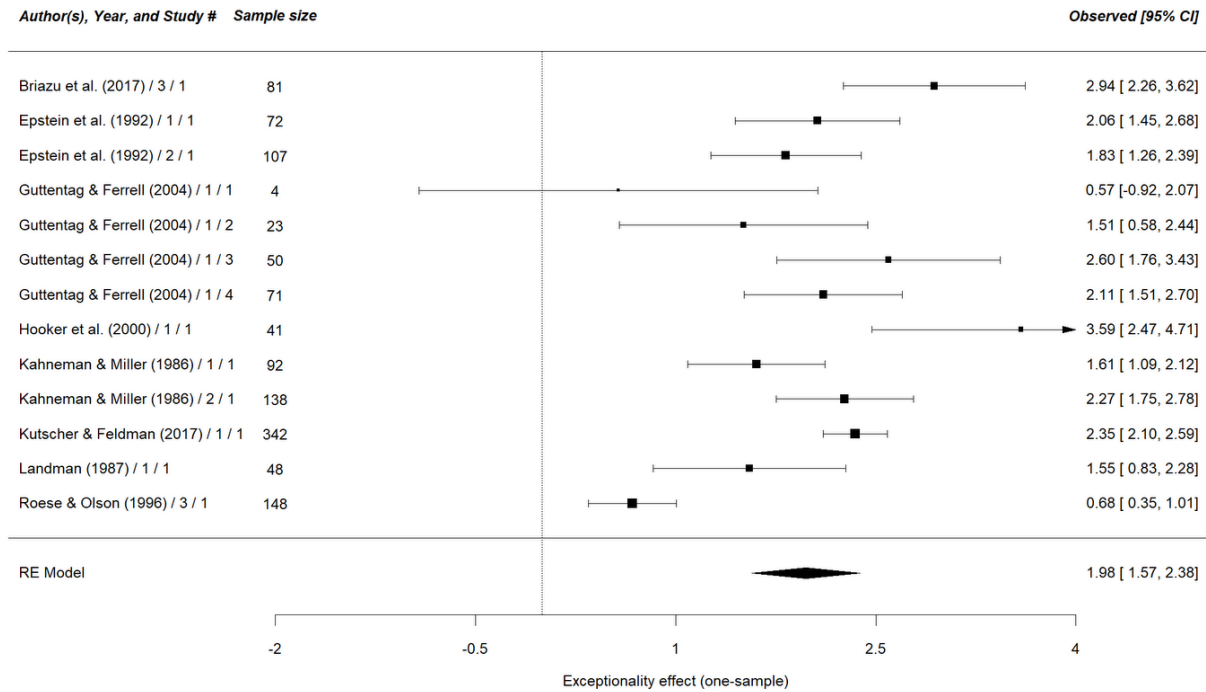
Meta-analysis forest plot for experimental studies

Figure 3

Trim-and-Fill funnel plot for experimental studies

Note. The eight missing studies are shown in black. We used the Trim and Fill method to see studies on the left with a random model, with the addition of the Egger regression shown as the red line. The Trim-and-Fill result should be taken with precaution, given the high heterogeneity found in our analysis.

Figure 4

Meta-analysis forest plot for one-sample studies

Exceptionality effect meta-analysis: Supplementary Materials

Open Science.....	2
Key decisions.....	2
Comparison of process versus pre-registration	3
Clarification regarding results	4
One-sample comparison studies: Counterfactuals	4
Funnel plot for one-sample studies.....	4
Pre-registration	4
Goals and research questions	4
Goal statement	4
Research questions	5
Hypotheses.....	5
Main hypothesis	5
Moderators.....	5
Exploratory moderators	7
Methods	8
Design.....	9
Search Strategy	9
Eligibility criteria	9
Inclusion criteria	10
Exclusion criteria.....	10
Procedure for studies selection.....	10
Data extraction (coding)	10
Analysis plan.....	11
Confirmatory analyses	11
Exploratory analyses	12
Conflicts of Interest	12
References.....	13
Contacting authors	15

Open Science

Data and code are shared on the Open Science Framework: <https://osf.io/542c7/>

Key decisions

We took several decisions beyond the pre-registration.

DV measures included in meta-analysis:

- (1) regret and general negative affect (being upset, foolish or disappointed)
- (2) number of counterfactuals (how likely someone wants change a certain forgone event) and avoidability
- (3) self-blame, victim responsibility and victim fault
- (4) victim compensation
- (5) offender punishment (e.g. fine, prison sentence)

Excluded DV measures:

- measures that were not explicitly related to regret
- luck
- victim positivity
- victim sympathy
- victim negligence

Comparison of process versus pre-registration

We note several deviations from the pre-registration protocol:

We reported Hedge's g rather than the pre-registered Cohen's d , to address reviewer comments we received on a pre-print. The effects and results were quite similar, yet Hedge's g is considered more accurate as it takes into account sample size.

Clarification regarding results

One-sample comparison studies: Counterfactuals

In the main manuscript, we indicated that counterfactuals had an effect of $g = 1.81$ $[-0.43, 4.05]$. This effect is mainly due to two factors: the number of studies for this effect is $k = 2$, and the methods we used to convert the effect size are different.

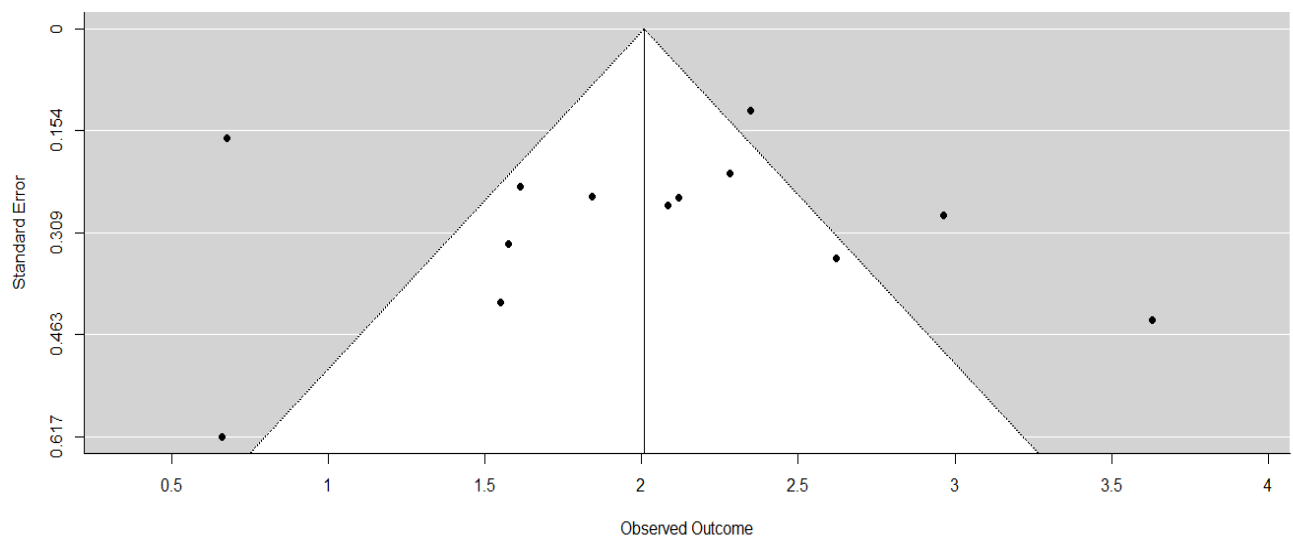
For the Briazu et al. (2017) study, participants had to compare two protagonists, and indicate who has more counterfactuals. 77 indicated that the protagonist in the exceptional condition has more counterfactuals, whereas only 4 indicated that the protagonist in the normal condition has more counterfactuals. For the Roes & Olson (1996) study, participants were asked how much counterfactual will the protagonist have. They recorded more counterfactuals in the exceptional ($M=1.17$) than routine ($M=.53$) antecedent condition, $F(1,144)=17.47$, $p=.001$.

One study measured a comparison between numbers of participant, the other measured a mean. This difference, in the addition of the $k=2$, can explain why we find a huge confidence interval including the null for this particular type of study.

Funnel plot for one-sample studies

We did not find enough one-sample studies to make accurate predictions about an average effect size, and we decided to remove the funnel plot to supplementary.

Meta-analysis funnel plot for one-sample studies



Pre-registration

We pre-registered the meta-analysis on July, 2017 on the Open Science Framework and data collection began after that.

Goals and research questions

Goal statement

In this meta-analysis, we aim to examine the impact of past-behavior normality on regret (confidence intervals do not include the null), determine overall effect size, and explore potential factors that moderate the effect.

Research questions

1. Do people associate stronger regret with negative outcomes that are a result of exceptional behavior in comparison to negative outcomes that are a result of routine behavior?
2. What is the overall effect size for the bias?
3. What are factors affecting the bias?

Hypotheses

Main hypothesis

The main hypothesis for the meta-analysis main effect:

Hypothesis: People tend to associate stronger regret with a negative outcome that is a result of exceptional behavior in comparison to routine behavior.

Moderators

Controllable vs. uncontrollable Events

Exceptions might not always be the result of personal preference. There can be multiple reasons why people make exceptions. Some people intentionally deviate from their everyday routines to enrich their lives with excitement and make new discoveries. But sometimes people are also forced to deviate from a certain routine by non-controllable circumstances. For example, the study of Miller & McFarland (1986) uses a scenario describing a man that always goes grocery shopping at the same supermarket. One day his regular shop is closed because of ongoing construction work, and he needs to visit another supermarket. In a second experimental group, the victim experiences the same fate after intentionally visiting another store. We argue that uncontrollable exceptions produce less regret. According to Decision Justifiability Theory (Connolly & Zeelenberg, 2002), regret is a consequence of self-blame over a decision that was hard to justify. Is someone forced to deviate from a routine the decision is easy to justify. The protagonist has no reason for self-blame because there simply is no other choice to make, so that perceived regret will be weaker in comparison to controllable exception. Additionally, an alternative version of the past in which one would do the normal thing might become less available because that option disappears from the set of choices. In line with this argumentation, studies report a tendency to mutate controllable events more likely than uncontrollable ones (Giroto, Legrenzi, & Rizzo, 1991; N'gbala & Branscombe, 1995). Taken together, the perception of regret is, therefore, likely to differ depending on how much control someone had over the exceptional behavior, such that controllable events elicit stronger regret.

The focal behavior will be coded for controllability (0 – not controllable, 1 – controllable, 99 – no indication of controllability or unable to deduce controllability).

Hypothesis: The impact of past-behavior routine-exception asymmetry on regret will be stronger when the behavior is controlled compared to when the behavior is uncontrolled.

Measure of the dependent variable: regret, vs. counterfactuals

We identified two common ways of assessing the impact of bias in exception-routine asymmetries, one referring to an emotional reaction – perceived regret, and one referring to mental simulation - counterfactual thoughts.

Several studies capture perceived regret, an emotion. After participants read a scenario, they are asked to judge the emotional reaction of the protagonist in that story. They, for example, have to compare the emotional states of two persons, as in the hitchhiker scenario – “Who do you expect to experience greater regret over the episode?” (Kahneman & Miller, 1986) or the airport scenario “Who is more upset?” (Tversky & Kahneman, 1982).

Other studies measure the number of produced counterfactual alternatives. For example, participants are presented with a scenario and are then asked which of the event they would like to mutate to make the negative event undone they are asked to rank certain events that they would like to undo. Common instructions are “list six things that could have been different to have changed the outcome” (Gavanski & Wells, 1989).

This moderator is exploratory. We do not yet have a specific hypothesis regarding the impact of the comparison between regret and counterfactuals, on the effect size of the main effect.

The dependent variable will be coded by type (1 - regret, 2- counterfactuals, 3 - others).

Exceptional vs. normal outcome

The first experimental scenarios that assessed mental simulation mainly used exceptional outcomes. Protagonist became a victim of rare incidents like a gunshot in a supermarket (Miller & McFarland, 1986), a car accident, a robbery by a hitchhiker (Kahneman & Miller, 1986) or a flight miss (Tversky & Kahneman, 1982). The first study that used normal and exceptional outcomes was Gavanski & Wells (1989) describing a story in which a good student receives either a high grade (normal) or a low grade (exception) after exceptional as well as routine events. They found that participants mainly mutated exceptional behavior in the direction of normality to undo exceptional outcomes, meaning that participants tried to mentally undo a bad grade of a good student after an exceptional event. These findings indicate that the exceptionality of the outcome has a considerable impact on the production of counterfactual thoughts and regret in the way that exceptional outcome is more likely to be altered as normal ones.

Outcome will be coded for exceptionality (0 - normal outcome, 1 - exceptional outcome, 99 – no indication of outcome exception or unable to deduce outcome exception or the difference from behavior exception).

Hypothesis: The impact of past-behavior routine-exception asymmetry on regret will be stronger when the outcome of the behavior is an exception compared to when it is a normal outcome.

Experimental design

N'gbala and Branscombe (1997) argued that the experimental design could serve as a potential moderator for norm related biases, and did not find an action-effect when they used between-person scenarios. Possibly, the effect is stronger when asymmetries are contrasted or compared by the same person.

(Coding: 0 – within person design, 1 – between person design)

Hypothesis: The impact of past-behavior routine-exception asymmetry on regret will be stronger when the experiment is using a within-subject design compared to when the experiment is using a between-subject design.

Routine strength

What makes a past behavior become routine? the issue with norm theory is that the definition for routine is flexible, there are no clear guidelines for what makes a routine become a routine. Scenarios vary in how they describe a routine. Some scenarios use the words “always” (Guttentag & Ferrell, 2004; Macrae et al., 1993) or “each day” (Turley et al., 1995), some refer to “often” (Hooker et al. 2000) or “regularly” (Macrae, 1992). Possibly, the stronger the routine or the stronger the word or the frequency used to describe a routine, the stronger the exception-routine effect would be.

Coding: 0 –medium frequency (e.g., several times, occasionally, etc.), 1 – high frequency (e.g., often, regularly, most times, habit, etc.), 2 – certainty (always, never, each day, etc.)

Hypothesis: The impact of past-behavior routine-exception asymmetry on regret will be stronger the stronger the routine.

Exploratory moderators

Action vs. Inaction

The action-effect (Kahneman & Tversky, (1982), is a phenomenon in which negative outcomes that are caused by an action (commission) result in stronger feelings of regret than negative outcomes that are a result of an inaction (omission). Norm theory (Kahneman & Miller, (1986) suggested that inaction can generally be seen as more normal in society than action. Action, therefore, can be considered a rather abnormal, leading to more counterfactual thoughts and therefore higher regret.

We first expect that there will be an effect regardless of action-inaction, and further expect that the effect of exception-routine asymmetry in past behavior would be even stronger when exception is aligned with taking action.

The focal behavior will be coded regarding whether the exception was aligned with taking action or not (0 – exception NOT aligned with action, 1 – exception aligned with action, 99-undetermined).

Hypothesis action-inaction #1: The impact of past-behavior routine-exception asymmetry on regret will be different from null (null not included in confidence intervals) even when exception not aligned with was action.

Competing hypothesis action-inaction #2a: The impact of past-behavior routine-exception asymmetry on regret will be stronger when exception behavior is

aligned with action and routine with inaction compared to when the exception behavior is aligned with inaction and routine with action.

Competing hypothesis action-inaction #2a: The impact of past-behavior routine-exception asymmetry on regret will not be stronger when exception behavior is aligned with action and routine with inaction compared to when the exception behavior is aligned with inaction and routine with action (criteria 1: z-test comparisons not significant; criteria 2: 95% confidence intervals overlapping)

Status quo

The status quo bias (Samuelson & Zeckhauser, (1988) described a phenomenon in which people favor a previously chosen or set option over a change to other alternatives. When faced with a decision people tend to stick with the status quo option. For example, Samuelson and Zeckhauser (1988) demonstrated that people who reconsider their financial investments prefer sticking with funds they already hold rather than investing in a new alternative, even when the change offered more attractive and profitable prospects.

Routine behavior might be confused for the status quo, yet there are important differences between routine behavior and status quo. In both cases, there is a common and known option that is associated with a positive emotional valence. However, status quo is about sticking to a previously chosen alternative or a default option. A status quo situation involves a previous decision favoring a specific option, and people are then likely to stay with that option when considering the decision. Routine represents a behavior that has been regularly repeated and has over time become the norm, such as driving home on the same route every day or daily shopping at the same supermarket.

We, therefore, expected that the experience of regret would be stronger when the protagonist has not only deviated from a routine but also from a status quo.

The decision situation will be coded regarding whether routine-exception were aligned with status-quo versus non status-quo (0 – not aligned, routine is different from the status quo, 1 – aligned, routine is the same as the status quo, 99 – no indication of status-quo or unable to deduce status-quo).

Hypothesis status-quo #1: The impact of past-behavior routine-exception asymmetry on regret will be different from null (null not included in confidence intervals) even when there is no alignment with the status quo.

Competing hypothesis status-quo #2a: The impact of past-behavior routine-exception asymmetry on regret will be stronger the routine behavior is aligned with the status quo compared to when the routine and status-quo are not aligned.

Competing hypothesis status-quo #2a: The impact of past-behavior routine-exception asymmetry on regret will not be affected by whether the routine behavior is aligned with the status quo compared to when the routine and status-quo are not aligned (criteria 1: z-test comparisons not significant; criteria 2: 95% confidence intervals overlapping)

Methods

Design

- Independent variable:
 - Past behavior normality: routine vs. exceptional behavior
- Dependent variables:
 - Perceived regret
 - Counterfactuals
- Moderators (see moderators and coding in section A above).

Search Strategy

- Database: Google Scholar (for suitability for meta-analyses see Walters, 2007; Gehanno, Rollin, & Darmoni, 2013).
- 15 combinations of the following search terms were used to search the database systematically. The exact combinations can be found in Appendix A.
 - General: norm theory, Miller, Kahneman
 - Bias names: normality bias, exceptional-routine effect, nonconformity effect, exceptionality effect
 - IV related: past behavior, past behavior norm, past behavior normality, abnormal behavior, intrapersonal, consistency, routine-violation, behavioral standard, behavioral standard, past action, exceptional events, routine events, routine, typical, atypical, unusual antecedent, usual antecedent, exceptional antecedent, repeated buying, repeated action, switching, self-blame regret, repurchase, prior usage
(note: search for both behavior and behaviour)
 - DV related: regret, upset, counterfactual, perceived regret, anticipated regret
- A scan of reference sections of found articles
- Search for “related articles” and “cited by” Google Scholar options of the identified articles
- Contacting authors of identified articles to ensure full coverage and maximize access to unpublished data and/or manuscripts
- Abstracts, tables and methods sections will be scanned to identify the relevance of a source.

Eligibility criteria

- Past behavior normality exception-routine asymmetry (IV) and measures of either counterfactuals or regret (DV)

Inclusion criteria

- Experimental designs only
- IV: Clear contrast between exception and routine
- DV: regret, counterfactuals, victim compensation, disappointment, negative affect, and self-blame (representing 80-90% of the identified literature);
 - o Note: Inclusion of studies using other regret-related measures such as responsibility, avoid ability, sympathy with the victim, and harm will be considered for inclusion during coding, but in any case, we will report with and without these additional studies.

Exclusion criteria

- DV related:
 - a. (please see note in inclusion above)
Exclusion of studies using other regret-related measures such as responsibility, avoid ability, sympathy with the victim and harm will be considered.
- IV related:
 - a. Studies that only manipulate general social normality or expectation normality.
 - b. Studies that only measure a deviation from a default option (status quo bias) which is only performed once and not routinely.
- Missing statistics are not reported: Studies which do not report crucial measures such as mean or standard needed for the calculation of the effect size deviation will be excluded from the sample.
- Correlational designs

Procedure for studies selection

Studies collected through the database searches will be assessed for their eligibility based on their titles, abstract and content. One researcher will determine the adequacy of the study for the meta-analysis and a second researcher will do the verification of the results. All the decisions to exclude a study will be documented with reasons.

All decisions on inclusion and exclusion will be documented in any case.

Data extraction (coding)

- A coding sheet will be prepared and pre-tested
- The coding sheet and code book are attached
- The coding process for the pretests will be completed by two coders to ensure a high inter-rater-reliability. Gaps identified will be documented, and decisions will be reported in detail.

- Once pre-test is completed, one coder will code all studies, the second coder will verify coding.

Analysis plan

We will use R and the metafor package for the statistical analyses. Given the range of different types of studies and experimental designs, we expect heterogeneity in the sample to be relatively high. Therefore, a random effects model will be used.

All effect sizes will be converted to Cohen's d and standardized to allow for a comparison. Split conditions due to moderators in the original studies will be collapsed to allow for a comparison of the main IV.

Whenever available, we will collect standardized effect sizes directly from authors of original papers. We will check for the accuracy of these analyses based on provided information and details. If unavailable we will use either descriptive statistics or inferential statistics to re-compute standardized effect sizes.

All conversions and coding decisions will be documented, and the original text will be included in the coding sheet to allow for reproducibility.

Forest plots presenting the effect size of each study will be produced. A meta-analysis will examine the overall main-effect, a meta-regression will be conducted to examine the impact of the described moderators.

Statistical heterogeneity will be determined using the τ^2 test and quantified using I^2 , which represents the percentage of total variation in a set of studies that is actually due to heterogeneity. (Higgins, Thompson, Deeks, & Altman, 2003). This global meta-analysis will yield a point estimate, confidence interval, and p-value, along with statistics for heterogeneity, assessed using the Q-statistics, and the I^2 statistic. If there is indeed significant heterogeneity, we will explore potential moderators.

We will report an analysis for the presence of publication bias, including funnel plots and statistical tests for publication bias (minimum: publication status as a moderator, compare effects for only published findings) and asymmetry (minimum: trim and fill, rank test, Egger's unweighted regression symmetry test).

We will also conduct a p-curve (Simonsohn, Nelson, & Simmons, 2014; Simmons, & Simonsohn, 2017) and a [p-uniform test](#) (van Aert & van Assen, 2017).

We aim to share all coding and R code with reviewers and the academic community using the Open Science Framework.

Confirmatory analyses

We will test for the hypotheses detailed in section A "Hypotheses" using a random-effects meta model.

We plan apriori to also conduct meta-analyses on subsets of the data, in particular, we will split the data by study design and IV/DV types.

Exploratory analyses

The coding sheet includes many other collected variables. We expect that will conduct additional exploratory on some of these variables, but those will be considered exploratory.

We also expect that additional hypotheses and possible coding moderators will be identified as we examine the papers and collected studies.

In both cases, we will explicitly declare these analyses as exploratory.

Answer the following final questions:

Has data collection begun for this project?

- ☒ ~~No, data collection has not begun~~
- ☐ **Yes**, data collection is underway or complete

If data collection has begun, have you looked at the data?

- ☒ ~~Yes~~
- ☐ **No**

The (estimated) start and end dates for this project are (optional):

Any additional comments before I pre-register this project (optional):

Conflicts of Interest

There are no conflicts of interest to report.

References

- AC Del Re (2013). compute.es: Compute Effect Sizes. R package version 0.2-2. URL <https://cran.r-project.org/package=compute.es>
- A. C. Del Re & W. T. Hoyt (2014). MAd: Meta-Analysis with Mean Differences. R package version 0.8-2. URL <https://cran.r-project.org/package=MAd>
- Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science*, 11, 212–216.
- Fan, F., Y. (2017). powerAnalysis: Power Analysis in Experimental Design. R package version 0.2.1. <https://CRAN.R-project.org/package=powerAnalysis>
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314–325.
- Giroto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78, 111–133.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136–153.
- Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246, 160–173.
- Kelley, K. (2019). MBESS: The MBESS R Package. R package version 4.6.0. <https://CRAN.R-project.org/package=MBESS>
- Kossmeier, M., Tran, U., S., & Voracek, M. (2020). metaviz: Forest Plots, Funnel Plots, and Visual Funnel Plot Inference for Meta-Analysis. R package version 0.3.1. <https://CRAN.R-project.org/package=metaviz>
- Lüdtke, D. (2019). esc: Effect Size Computation for Meta Analysis (Version 0.5.1). doi: 10.5281/zenodo.1249218. <https://CRAN.R-project.org/package=esc>.

- Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, 12, 513–519.
- N'gbala, A., & Branscombe, N. R. (1995). Mental simulation and causal attribution: When simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, 31, 139–162.
- Revelle, W. (2019) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 1.9.12.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1, 7–59.
- Tversky, A., & Kahneman, D. (1982). The simulation heuristic. *Judgment under Uncertainty: Heuristics and Biases*, 201–208.
- Van Aert, R., C., M. (2019). puniform: Meta-Analysis Methods Correcting for Publication Bias. R package version 0.2.1. <https://CRAN.R-project.org/package=puniform>
- Van Lissa, C., J. (2020). metaforest: Exploring Heterogeneity in Meta-Analysis using Random Forests. R package version 0.1.3. <https://CRAN.R-project.org/package=metaforest>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. URL: <https://www.jstatsoft.org/v36/i03/>
- Walters, W. H. (2007). Google Scholar coverage of a multidisciplinary field. *Information processing & management*, 43, 1121-1132.

Contacting authors

We contacted authors with the following call:

Dear colleagues,

We are conducting a meta-analysis on how past behavior normality impacts perceived counterfactual regret after a negative outcome (e.g., hitchhiker-scenario by Kahneman & Miller 1986 or Miller & McFarland, 1986). The bias is also known as the exceptional-routine effect, exceptionality effect or normality bias. Some scholars also refer to past behavior normality with terms as behavioral consistency or intrapersonal norms. We want to assess how a deviation from such a norm impacts counterfactual regret after a negative outcome. We have completed an initial exhaustive search of the literature and have identified you as an author who has published work on the topic, and so we are contacting you to ask for your unpublished manuscripts and data to be included in the meta-analysis.

We would appreciate references to your published data to make sure we have included it in our meta. More importantly, we are especially interested in any relevant unpublished manuscripts or data that cannot be found using regular literature search.

If you have unfinished or unpublished manuscripts, we would appreciate a copy. Alternatively, for unpublished manuscripts and/or data, the information we require for inclusion is:

- A description of the manipulation and general description of the experimental conditions.
- For each experimental condition:
 - Brief description of the condition
 - Sample size
 - Mean and standard deviation for each of the dependent variables
 - Brief description of the measures/scales were used for the dependent variables, and internal reliabilities if available/relevant.
- Sample characteristics, such as: overall sample size, country, sample type (students, MTurk, general population, etc.), mean age.
- Reference to be used when citing this data.

If you only have raw data that has not yet been analyzed, then we would be happy to help analyze it for inclusion. In such a case, please send us the dataset and a description of the key variables described above relevant for the analysis.

Please send all relevant information and/or data to Lucas Kutscher at email:

l.kutscher@student.maastrichtuniversity.nl

If you are interested in more information, we will be happy to answer any further questions.

Additionally, you can follow updates and read further information on our ResearchGate project:

<https://www.researchgate.net/project/The-impact-of-past-behavior-norms-on-regret-Pre-registered-replication-and-meta-analytic-review>

Best regards,

Gilad Feldman & Lucas Kutscher

Department of work and social psychology
Maastricht University, Netherlands