

The impact of past behaviour normality on regret: replication and extension of three experiments of the exceptionality effect

Lucas Kutscher & Gilad Feldman

To cite this article: Lucas Kutscher & Gilad Feldman (2019) The impact of past behaviour normality on regret: replication and extension of three experiments of the exceptionality effect, *Cognition and Emotion*, 33:5, 901-914, DOI: [10.1080/02699931.2018.1504747](https://doi.org/10.1080/02699931.2018.1504747)

To link to this article: <https://doi.org/10.1080/02699931.2018.1504747>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 10 Aug 2018.



[Submit your article to this journal](#)



Article views: 878




[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

The impact of past behaviour normality on regret: replication and extension of three experiments of the exceptionality effect

Lucas Kutscher^a and Gilad Feldman ^{a,b}

^aDepartment of Work and Social Psychology, Maastricht University, Maastricht, the Netherlands; ^bDepartment of Psychology, University of Hong Kong, Hong Kong SAR, China

ABSTRACT

Norm theory (Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136–153) described a tendency for people to associate stronger regret with a negative outcome when it is a result of an exception (abnormal behaviour) compared to when it is a result of routine (normal behaviour). In two pre-registered studies, we conducted a replication and extension of three classic experiments on past behaviour exception/routine contrasts ($N = 684$). We successfully replicated Kahneman and Miller's (1986) experiments with the classic hitchhiker-scenario (Part 1) and car accident-scenario (Part 2). In both cases, participants examined negative outcomes and tended to indicate a protagonist who deviated from own past behaviour as more regretful than another who followed routine. Pre-registered extensions also showed effects for ratings of social norms, negative affect, and perceived luck. We did not find support for the Miller, D. T., and McFarland, C. [(1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, 12, 513–519] experiment robbery scenario (Part 3) using a compensation measure, in that compensation to a victim of a robbery was not significantly different comparing exceptional and routine circumstances. However, a pre-registered extension showed that robbery under exceptional circumstances was perceived as more regretful than robbery under routine circumstances. We discuss implications for current and future research.

ARTICLE HISTORY

Received 25 September 2017
Revised 16 June 2018
Accepted 21 July 2018

KEYWORDS

Normality; norm theory; regret; exception routine; exceptionality effect; replication

Regret is a negative emotion that plays an important role in people's daily lives. Regret is experienced when perceiving that a different decision or circumstances could have led to a more positive or desirable result (Byrne, 2016; Roese & Epstude, 2017; Kahneman & Tversky, 1982).

In their influential work on norm theory, Kahneman and Miller (1986) highlighted normality as a core determinant of regret and counterfactual thought – imagined alternative versions of the past about what could have happened differently (Byrne, 2016; Miller, Turnbull, & McFarland, 1990; Roese, 1997). They summarised evidence suggesting that abnormal behaviour tends to elicit more counterfactual thought and

stronger feelings of regret. They discussed the concept of normality broadly, covering experimental studies referring to several factors that affect perceptions and/or construal of what is perceived as normal (Feldman & Albarracín, 2017; Feldman, 2018). Some of the experimental evidence they reviewed referred to past behaviour as the reference point to which behaviours are compared in evaluating whether a person's behaviour is normal or exceptional for the person. The past-behaviour normality effect was illustrated using a car-accident scenario thought experiment: Mr. White is described as driving home using the same route every day, and the one day he decided to take a different route he had an accident.

CONTACT Gilad Feldman  gfeldman@hku.hk

 Supplemental data for this article can be accessed here <https://doi.org/10.1080/02699931.2018.1504747>

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

A clear majority of participants (82%) indicated that Mr. White is more likely to be upset than someone who had a similar accident driving on his usual route. Recent literature (Byrne, 2016; Bear & Knobe, 2017; Dixon & Byrne, 2011; McEleney & Byrne, 2006) referred to this bias by using the term *exceptionality effect*, which we will also use in the present investigation, focusing on examining normality in reference to past-behaviour.

Kahneman and Miller's (1986) norm theory argued that events are mentally classified as normal or abnormal, and that abnormal events evoke stronger reactions since normal instances are more easily retrieved from memory, making it easier to elicit counterfactual alternatives to abnormal events. Higher availability of these mental simulations intensifies regret over a negative outcome. People's routine behaviour, for example, is easier to recall and is mentally more salient, making it easier to imagine routine counterfactual realities to an exception. Deviations from one's past actions are, therefore, more mutable and trigger more intense feelings of regret.

Impact of norm theory and the concept of normality

Norm theory has been an influential theory in the domains of cognition and emotion with important implications for behaviour across domains, and the concept of normality has served as a meaningful construct in understanding counterfactual thought and associated affect (Roese & Olson, 1997). Normality has been conceptualised as a key factor in both automatic and controlled processing, with both descriptive and prescriptive representations of normal, used to set reference points in making assessments of situations (Bear & Knobe, 2017). A two-stage model of counterfactual thinking (Roese, 1997; Roese & Hur, 1997) suggested that counterfactual thought is activated by negative affect, and normality then determines content to serve as a reference point, in aim of directing behaviour to return to a normal state (Buck & Miller, 1994; Miller, Taylor, & Buck, 1991; Miller et al., 1990; Wells, Taylor, & Turtle, 1987). Normality is also associated with causal interpretations and inferences (Gavanski & Wells, 1989; Roese & Olson, 1996). The concept of normality is functional in that it is predominantly beneficial for the person to refer to and rely on what is perceived as normal in terms of efficiency, social functioning, and survival (Roese, 1997). Using the functional perspective of

counterfactual thought (Epstude & Roese, 2008; Roese & Epstude, 2017) normality signifies the desired state, with affect and counterfactual thought signalling deviations from that state. This then allows for learning from such deviations, and then readjusting by guiding future behaviour through means of change in goals or motivation.

Choice of studies for replication

The aim of the present study is to conduct a direct replication of three classic experiments demonstrating the exceptionality effect, considered the earliest studies investigating the impact of past behaviour normality on regret. Parts 1 and 2 of this replication use the classic hitch-hiker and car-accident scenarios from Kahneman and Miller (1986). Part 3 is a replication of Experiment 1 from Miller and McFarland (1986) with the robbery-scenario – the study that meant to provide the first empirical test of norm theory.

Based on a meta-analysis of the exceptionality effect (Kutscher & Feldman, 2018) and to the best of our knowledge there are no documented attempts to conduct a direct replication of the three experiments. According to Google Scholar, at the time of writing, Kahneman and Miller's (1986) article has been cited over 3200 times and Miller and McFarland's (1986) article has been cited over 275 times with over 900 and 170 of those citations, respectively, including references to exception or routine and to regret or counterfactuals.¹ These indicate the very high impact of these studies over the field and the importance in revisiting and elaborating these effects.

Replicated studies and extensions

In Part 1 we aimed to replicate findings from Kahneman and Miller (1986) using the hitchhiker-scenario describing a simple comparison between two persons: Mr. Smith usually takes hitchhikers and Mr. Jones almost never does. They both offer a ride to a hitchhiker and get robbed.

In the original experiment, most participants (88% of 138 participants) believed that Mr. Jones – who acted abnormally – feels greater regret than Mr. Smith who acted as he usually did, thereby reflecting regret asymmetry in favour of exceptions in comparison to routines.

In addition to compared regret, the original setup also included a question asking who of the two is

more likely to be criticised for his behaviour (“who will be criticised most severely by others?”). This measure introduced another normality type, referring to norms within a certain group of people, such as societies, families, or organisations (social normality). A total of 77% of all participants rated Mr. Smith’s behaviour as more likely to be criticised severely. Research has shown the impact of social norms on perceptions of regret, with higher regret associated with acting abnormally with regards to social norms (Feldman, 2018; Feldman & Albarracín, 2017; Koehler & Prentice, 2003; Koonce, Miller, & Winchel, 2015).

In the hitchhiker-scenario, both protagonists behave abnormally, Mr. Jones with regards to past behaviour norms and Mr. Smith with regards to social norms. In this case, the exceptional behaviour of Mr. Jones’ behaviour is more salient thus past behaviour normality has the stronger impact on regret. Contrary to these findings there was a different study (Hur, Roese, & Namkoong, 2009) which found that social norms were stronger than past behaviour norms. The mixed evidence motivated us to extend the replication study with two additional measures of social norms. Finally, a measure of general negative affect was added, to assess the overall anticipated emotional reaction of the protagonists combining both norm types.

Part 2 of this replication tested Kahneman and Miller’s (1986) car accident-scenario. They reported that people associated stronger regret (82% of 92 participants) with a man who had an accident while driving home and deviating from his routine route, in comparison to a someone who had a similar accident while driving home on his normal route. Variations of this scenario were conceptually replicated and extended to different settings with relatively small samples (Briazu, Walsh, Deepröse, & Ganis, 2017: $N=102$; Epstein, Lipson, Holstein, & Huh, 1992: $N=83$; Hooker, Roese, & Park, 2000: $N=43$). Other studies describe more extreme incidents such as rape or robbery (Macrae, Milne, & Griffiths, 1993; Turley, Sanna, & Reiter, 1995; Wells et al., 1987) and using a variety measures. However, as far as we know all studies have used slight changes in the story description and measured regret with different approaches, so that this study is the first attempt to directly attempt a replication of the car accident-scenario by Kahneman and Miller (1986) in its original version.

The replication was extended with additional measures of perceived random chance and luck.

Miller, Turnbull, and McFarland (1989) showed that people evaluate an event to be more normal and generate weaker emotional reactions the easier it is to recall similar events. For example, people tend to be more suspicious of a manager’s intentions when a manager of a company with nine women and 90 men decides to promote a man, in comparison to a manager promoting a man in a company with only one woman and ten men. The random chance of promoting a woman is 10% in both cases, but people believed that promotion of a man candidate is more normal when the absolute number of paths through which the outcome can possibly occur was higher (nine paths in comparison to only one path).

We hypothesised that this principle would apply to exceptions and routines. A routine route is driven more often, and the absolute number of accidents that can potentially occur is, therefore, higher compared to an unusual route. Thus, people can more easily imagine an accident on a usual route in comparison to the unusual one because more paths of action could potentially lead to this event. The randomness and luck measures were added as exploratory measures to assess whether people’s estimations of probabilities moderate their perception of normality.

In Part 3 we conducted a replication of the robbery-scenario in Miller and McFarland’s (1986) Experiment 1. They tested the effect of past behaviour normality in an experiment comparing between three experimental groups. People were asked to evaluate the fate of a man who became the victim of a robbery occurring at a grocery store. As a manipulation of past behaviour participants were informed whether the man was shopping at his regular market or that he shopped in a store that he rarely visits. They found that people assigned higher victim compensation when the incident was a result of exceptional behaviour (reported as “ $t(162) = 2.17, p < .03$ ”; further calculated as $d = 0.36$ CIs [0.03; 0.68], see supplementary). We aimed for a direct replication and further extended the compensation measure by adding a measure of regret that is empirically closer to the original Kahneman and Miller (1986) paradigm.

Experiments 1–2

Pre-registration and open-science

There was an eight months gap between the two replication attempts. In each of the replication studies, we first pre-registered the experiment on the Open

Science Framework and data collection was launched later that day. Pre-registrations, power analyses, and all materials used in these experiments are available in the supplementary materials. These together with data and code were shared on the Open Science Framework (Experiment 1 – DOI 10.17605/OSF.IO/ZA7T2; Experiment 2 – DOI 10.17605/OSF.IO/P4RGD).

Differences between the two replications attempts are discussed at the end of the methods section.

Participants and general procedure

We recruited American Amazon Mechanical Turk (MTurk) participants online using TurkPrime.com (Litman, Robinson, & Abberbock, 2017).

There were 342 participants in the first replication experiment (182 males, 156 females, 4 unreported, $M_{age} = 38.28$, $SD_{age} = 11.55$), and 342 participants in the second replication experiment (127 males, 215 females, $M_{age} = 39.93$, $SD_{age} = 12.88$). We ensured that participants that took part in the first experiment could not take part in the second. In the second replication we also addressed MTurk possible non-naiveite (Chandler, Mueller, & Paolacci, 2014), by disallowing 5% of the top active workers, and setting a maximum for number of tasks performed on the platform (features available in Turkprime).

The overall experimental design included three parts. The first and the second parts of this experiment were a replication of experimental scenarios and data presented in the Kahneman and Miller (1986) review paper – the hitchhiker-scenario and the car accident-scenario. The third part was a replication of the robbery scenario from Miller and McFarland's (1986) Experiment 1.

Part 1: Hitchhiker-scenario

Participants evaluated a classic scenario, known as the hitchhiker-scenario by Kahneman and Miller (1986, p. 145) assessing the impact of past behaviour on the perception of regret following a negative outcome.

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.

The scenario was followed by three comprehension questions to make sure that the scenario was well understood by the participants – “who almost never takes hitch-hikers in his car?” (Mr. Jones / Mr. Smith),

“who frequently takes hitch-hikers in his car?” (Mr. Jones / Mr. Smith), “who got robbed?” (Mr. Jones / Mr. Smith / Both). The participants had to answer the comprehensions questions correctly before proceeding to the next page.

We first assessed regret to attempt a replication of the original experiment. In addition to the measure of regret, we extended the design by adding two measures going beyond the replication. The extension questions were presented on separate pages to eliminate any possible effects over the original regret measure. We examined descriptive and injunctive social norms. Descriptive norms are related to behaviour that is usually performed by others, whereas injunctive norms assess how acceptable a certain behaviour is in society (Cialdini, 2003). Based on Kahneman and Miller (1986) we expected the behaviour of the person regularly taking hitch-hikers to be evaluated as less common (descriptive norms) and more likely to be criticised by society (injunctive norms). The second extension to the original experiment was a measure of overall negative affect considering both social norms and past behaviour.

Measures

Regret. Regret was measured using the following question adapted from the original study – “who do you expect to experience greater regret over the episode? (Mr. Jones / Mr. Smith)”.

Norms. Participants answered two questions regarding perceived social norms, measuring descriptive norms – “whose behaviour do you think is more common in society? (Mr. Jones / Mr. Smith)” and injunctive norms – “whose behaviour do you think will be more criticised by others in society? (Mr. Jones / Mr. Smith)”.

Negative affect. The measure was implemented as an explanatory variable with no specific hypothesis – “contemplating your previous answers about this scenario and factoring in both Mr. Jones and Mr. Smith personal routines and your perceptions of social norms and possible social criticism, who do you think overall experienced more negative feelings about the decision to take a hitch-hiker that day?” (Mr. Jones / Mr. Smith).

Part 2: Car accident scenario

Participants were presented with the car accident-scenario introduced in Kahneman and Miller (1986, p. 145):

Mr. Adams was involved in an accident when driving home after work on his regular route.

Mr. White was involved in a similar accident when driving on a route that he only takes when he wants a change of scenery.

The scenario was followed by three comprehension questions to make sure that the story was well understood by the participants – “who was driving home after work on his regular route?” (Mr. Adams / Mr. White), “who was driving on a route that he only takes when he wants a change of scenery?” (Mr. Adams / Mr. White), “who was involved in an accident?” (Mr. Adams / Mr. White / Both).

We first assessed regret as a replication of the original experiment. We added two new measures to the experiment presented on separate pages after the replication to avoid interference with the original measures of regret. Relating to the study of Miller et al. (1989) measures of random chance and luck were implemented as exploratory variables.

Measures

Regret. Regret was measured using the following question adapted from the original study – “who is more upset over the accident?” (Mr. Adams / Mr. White).

Randomness. Participants rated perceived randomness – “Mr. Adams’ accident is just a random coincidence” and “Mr. White’s accident is just a random coincidence” (1 – *Strongly disagree*; 7 – *Strongly agree*).

Luck. Participants rated perceived luck – “which of the two do you think is less lucky?” (Adams / White).

Part 3: Robbery scenario

The last part of this study was a direct replication of Experiment 1 in Miller and McFarland (1986). The exact scenario used by the authors was not provided in the article and was therefore reconstructed. Participants were randomly and evenly assigned to three experimental conditions in a between-subject design. Participants read a scenario describing a victim of a robbery taking place at a grocery store. The scenarios only differed in normality of the circumstances for the protagonist leading up to the incident. In the first condition (routine behaviour) Mr. Paul visits the store he regularly visits. In the second condition (self-produced exception) the protagonist visits another store because he wants a change of pace. In the third condition (other-produced exception) Mr.

Paul is forced to visit another store because his regular store is closed due to ongoing construction.

Two convenience stores are located in Mr. Paul’s neighborhood. He frequents Store A more regularly than Store B.

[*Routine behavior condition:* Last night he visited Store A.

Self-produced exception condition: Last night he visited Store B because he wanted a change of pace.

Other-produced exception condition: Last night he visited Store B because Store A was temporarily closed for renovations.]

He walked in on a robbery occurring in the store. He lost the use of his right arm as a result of a gunshot wound.

In Experiment 1 the scenario was followed by three comprehension questions that participants had to answer correctly in order to proceed to the measures – “which convenience store does Mr. Paul visit frequently?” (Store A / Store B), “which convenience store did Mr. Paul visit last night?” (Store A / Store B), “did Mr. Paul lose the use of his right arm as a result of a gunshot wound?” (Yes / No). Experiment 2 included no comprehension questions. The measure of compensation was presented first, and the regret measure second.

Measures

Compensation. Participants were asked to decide on a compensation to the victim: “How much money should Mr. Paul receive in compensation for his loss?” (11-point scale: 0–0 \$ to 10–1,000,000 \$; typical award: 500,000 \$). In Experiment 2 the question varied slightly to address feedback received from the authors of the target article – “Mr. Paul seeks compensation for both the physical and psychological harm suffered. How much money should Mr. Paul receive in compensation?.” Also, since the responses in Experiment 1 converged around the “typical award” of 500,000\$ it is possible that this may have affected the results, and so Experiment 2 did not include an indication of what a typical reward is.

Regret. Participants rated perceived regret – “assume there was no compensation given to Mr. Paul. How much regret does he feel over the situation?” (1 – *no regret* to 5 – *very strong regret*). In Experiment 2 we clarified this even further to be about the actual decision made to visit the specific store and not about the unfortunate circumstances – “Assume there was no compensation

given to Mr. Paul. How much regret does he feel about visiting store [A/B]?"

Differences between the two replication attempts

The purpose of the second replication attempt was to address the failure to replicate findings in Experiment 1 Part 3 regarding the compensation dependent variable originally used in the target article (presented and discussed below). In the second replication attempt we only collected data for Part 3, thereby addressing possible concerns that exposure to questions in Part 1 and 2 somehow affected answers in Part 3. We also removed comprehension checks used in the first replication study in order to more closely mirror the original study design. Based on suggestion made by the original authors of the target article we also made it clear that compensation is about both physical and psychological harm. Finally, since the findings in the first replication were centred around the indicated "typical award (500,000 \$)", we removed the indication of a typical award in the second replication.

Deviations from the original studies

Together, the two experiments provide for a very close direct replication of the procedures by the original studies, with the main difference being the target sample.

Results

Part 1: Hitchhiker-scenario

Part 1 was only included in Experiment 1. All measures were analysed using a chi-square test to assess whether the distribution significantly deviated from a 50-50 distribution indicating a random choice with no preference to either option ($p = .5$). The results are provided in [Table 1](#) and visualised in [Figure 1](#).

Regret

On basis of the original study, we expected participants to assign stronger feelings of regret to Mr. Jones, who almost never takes hitchhikers. Most participants associated higher regret with Mr. Jones (315, 92%) rather than Mr. Smith (27, 8%; $\chi^2(1, N = 342) = 242.53, p < .001$).

Our findings replicated the original study in which most participants rated Mr. Jones as experiencing stronger regret.

Social norms

Going beyond the replication, we expected participants to rate Mr. Jones' behaviour (never taking hitchhikers) as more normal in society than Mr. Smith's (often taking hitchhikers). Most participants rated Mr. Jones's behaviour as more common (326, 95%), compared to that of Mr. Smith (16, 5%; $\chi^2(1, N = 342) = 280.99, p < .001$). Mr. Smith was evaluated as more likely to be criticised by others (310; 91%) than Mr. Jones (32, 9%; $\chi^2(1, N = 342) = 225.98, p < .001$).

Our findings supported our hypotheses. Most people assessed Mr. Jones' behaviour (never taking hitchhikers) as more common in society, while Mr. Smith's behaviour (taking hitchhikers) was evaluated as more likely to be criticised by others.

Negative affect

The second extension of the replication assessed the general negative affect combining past behaviour and social norms. Most participants rated Mr. Jones as feeling worse than Mr. Smith (Jones: 317, 93%; Smith: 25, 7%; $\chi^2(1, N = 342) = 249.31, p < .001$).

Part 2: Car accident-scenario

Part 2 was only included in Experiment 1. Regret and luck were analysed using a chi-square test comparing proportions to a 50-50 random choice split ($p = .5$). Randomness means were compared by performing a dependent sample *t*-test. The results are provided in [Table 2](#) and visualised in [Figure 2](#).

Regret

We expected that participants would assign higher levels of regret to Mr. White who had an accident on an exceptional route home. Replicating the original results, most participants associated stronger regret with the exception (277, 81%) than with routine behaviour (65, 19%; $\chi^2(1, N = 342) = 131.42, p < .001, d = 1.58$).

Randomness

Participants rated Mr. Adams' (exception, $M = 5.74, SD = 1.23$) fate as slightly more likely to be a random coincidence than the fate of Mr. White (routine, $M = 5.67, SD = 1.23$). A dependent samples *t*-test indicated

Table 1. Part 1 (hitchhiker): Counts and proportions for perceived regret, social norms, and negative affect.

	Regret		Social norm (injunctive)		Social norms (descriptive)		Negative affect	
	Counts	%	Counts	%	Counts	%	Counts	%
Routine Smith	27	7.9	16	4.7	310	90.6	25	7.3
Exception Jones	315	92.1	326	95.3	32	9.4	317	92.7

that the difference was not statistically significant, $t(339) = 1.50$, $p = .134$, $d = 0.06$.

Luck

Participants perceived Mr. White (exception) to be less lucky than Mr. Adams (routine), with 114 participants (33%) considered Mr. Adams as less lucky, whereas 228 (66%) felt that Mr. White was less lucky, $\chi^2(1, N = 342) = 38.00$, $p < .001$, $d = 0.71$).

Part 3: Robbery scenario

Part 3 was included in both Experiments 1 and 2. In the original study, the authors found no significant difference between the self-induced exception condition and the other-induced exception condition and they, therefore, chose to collapse the two conditions into a single exception condition. We first provide results using the original analysis with collapsed conditions and a t -test comparing between routine and exception. Additionally, we analysed the differences between all three conditions using an ANOVA and a t -test contrasting condition pairs. Results are provided in Table 3 and plotted on Figures 3 and 4.

Compensation

An independent samples t -test comparing the two conditions revealed no significant difference between compensation in the joint exception and the routine condition (Exp1: $t(246) = 0.79$, $p = .428$, $d = 0.09$; Exp2: $t(242.59) = 0.53$, $p = .596$, $d = 0.05$). There were also no significant differences in compensation to a victim between all three conditions in either Experiment 1 ($F(2, 339) = 0.74$, $p = 0.493$; all t -test contrasts $|t| < 1.19$, $p > .236$, Cohen $|d| < 0.16$) or

Table 2. Part 2 (car accident): Counts and proportions for perceived regret and luck.

	Regret		Luck	
	Counts	%	Counts	%
Routine Adams	65	19.0	114	33.3
Exception White	277	81.0	228	66.7

Experiment 2 ($F(2, 339) = 0.17$, $p = 0.845$; all t -test contrasts $|t| < 0.58$, $p > .564$, Cohen $|d| < 0.08$).

In both experiments a Shapiro–Wilk test of normality revealed that the distribution was not normal ($p < .001$), and so we also ran a series of non-parametric Mann–Whitney U tests, which again showed no significant differences between the conditions (Exp1: $p > .207$; Exp2: $p > .657$). We note that the changes we made in Experiment 2 to remove the indicator for a “typical award” of 500,000 mid-point resulted in a much wider distribution yet very similar results in terms of the contrasts between the conditions.

We conclude that we failed to replicate the findings of the original compensation measure by Miller and McFarland (1986).

Regret

With collapsed exceptional conditions an independent samples t -test revealed significantly stronger regret in the exceptional conditions in comparison to the routine condition (Exp1: $t(184) = 3.99$, $p < .001$, $d = 0.50$; Exp2: $t(178.42) = 4.44$, $p < .001$, $d = 0.56$). Comparing the three conditions, participants rated stronger regret for self-produced exception and other-produced exception than to routine in both Experiment 1 (self-produced exception versus routine: $t(204) = 4.53$, $p < .001$, $d = 0.60$; other-produced exception versus routine: $t(219) = 2.71$, $p < .001$,

Table 3. Part 3: Sample size, means, and standard deviations for victim compensation and regret.

Condition	N	Compensation		Regret	
		M	SD	M	SD
Experiment 1					
Routine	114	5.34	2.37	3.95	1.12
Self-produced exception	115	5.72	2.46	4.53	0.80
Other-produced exception	113	5.40	2.76	4.32	0.94
Exception (combined)	228	5.57	2.61	4.43	0.87
Experiment 2					
Routine	116	4.74	3.40	3.96	1.19
Self-produced exception	110	4.89	3.50	4.69	0.67
Other-produced exception	116	5.01	3.66	4.34	0.98
Exception (combined)	226	4.95	3.57	4.50	0.86

Note: Scales: Regret – 1 to 5; Compensation – 1 to 10.

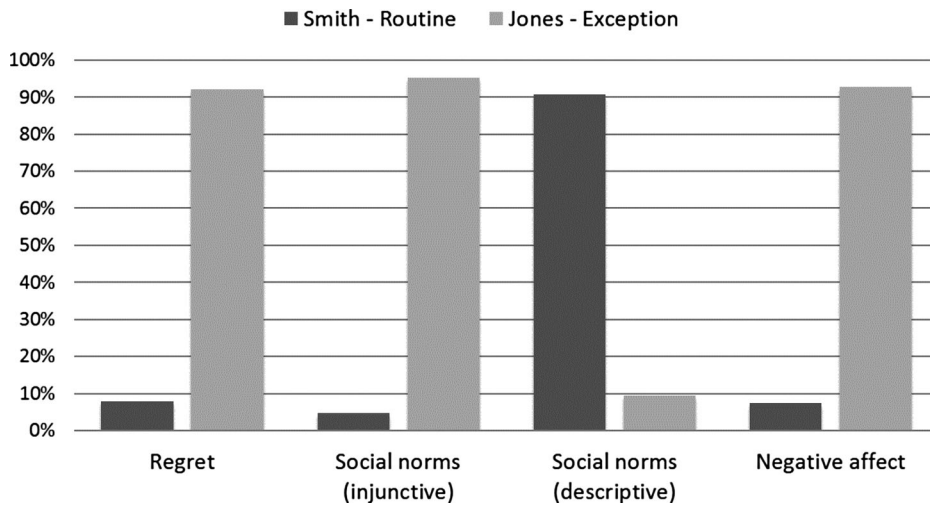


Figure 1. Part 1: Proportions for perceived regret, injunctive social norms, descriptive social norms, and negative affect.

$d = 0.36$; $F(2, 339) = 10.80$, $p < .001$) and Experiment 2 (self-produced exception versus routine: $t(183.77) = 5.74$, $p < .001$, $d = 0.75$; other-produced exception versus routine: $t(221.66) = 2.65$, $p = .009$, $d = 0.35$; $F(2, 339) = 16.02$, $p < .001$). The differences between the two exception conditions were weak and not significant in Experiment 1 ($t(219) = 1.84$, $p = .068$, $d = 0.24$) yet slightly stronger and significant in Experiment 2 ($t(204.86) = 3.19$, $p = .001$, $d = 0.42$).

The regret findings supported our pre-registered hypotheses.

Comparing target article and replications' effect size

Comparing the replication findings to the target article, the findings for the two scenarios in Kahneman and Miller (1986) were very similar (hitchhiker scenario: original 88%/12%, replication 92%/8%; car accident scenario: original 82%/18%; replication 81%/19%).

The effect size for Miller and McFarland's (1986) supermarket scenario for the contrast between the routine and exception conditions on the compensation

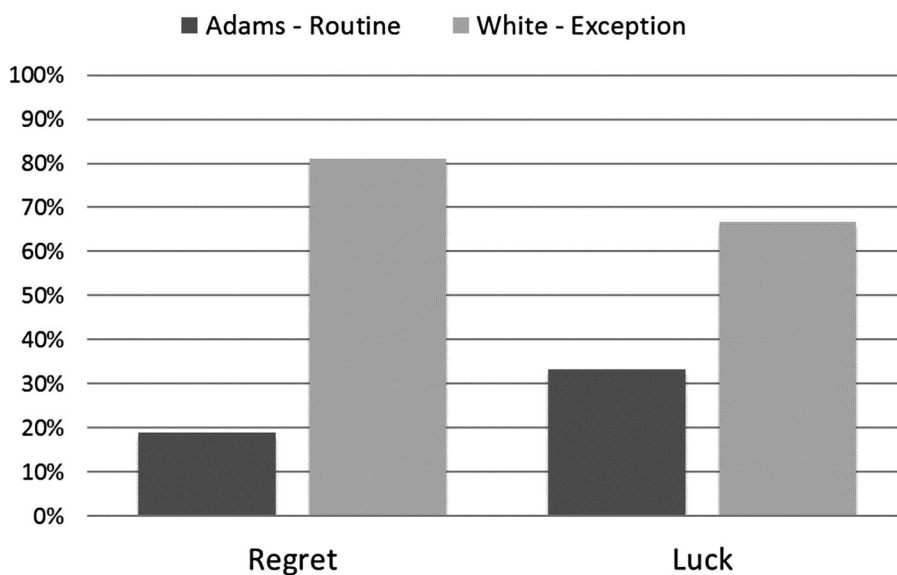


Figure 2. Part 2: Proportions for perceived regret and luck.

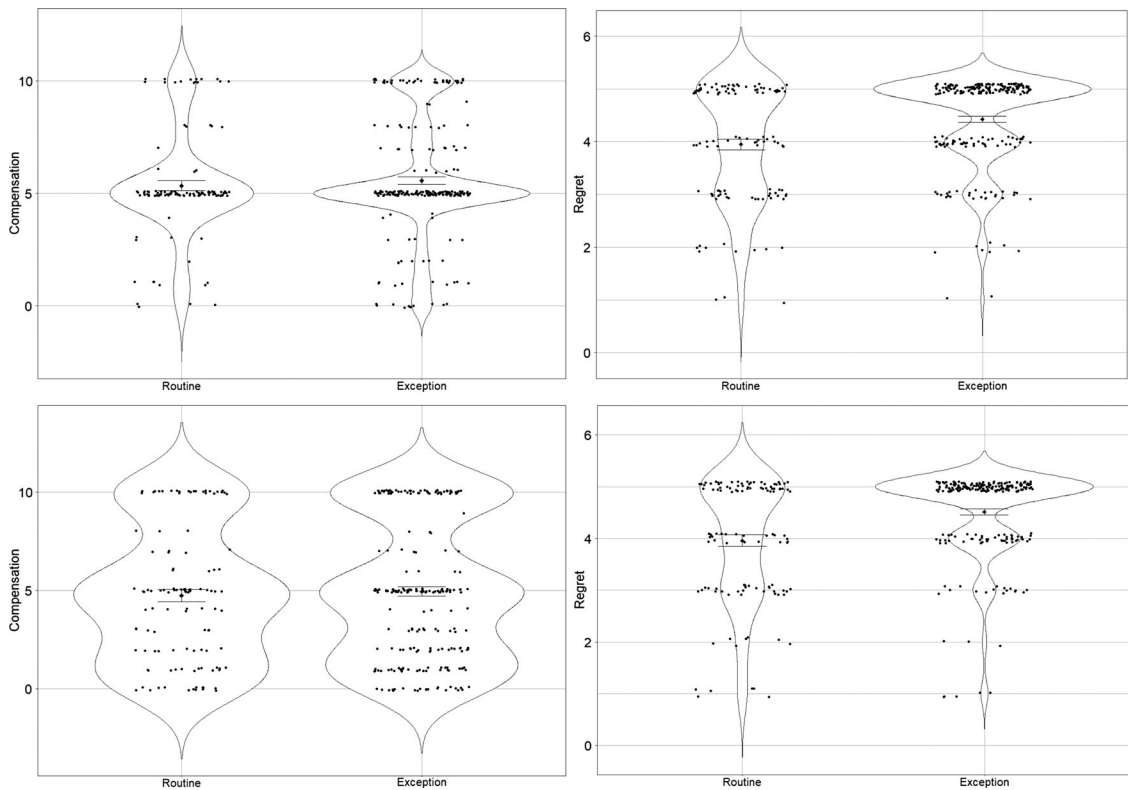


Figure 3. Part 3: Victim compensation and regret (with exception conditions combined). The top two plots are for Experiment 1, the bottom two plots are for Experiment 2. Error bars represent standard error. Scales: Regret – 1 to 5; Compensation – 1 to 10.

measure was $d = 0.36$ [0.03, 0.68] (routine: $n = 58$, $M = 4.52$; self-produced exception: $n = 48$, $M = 5.37$; other-produced exception: $n = 57$, $M = 5.37$; $t(162) = 2.17$, $p < .03$). In both experiments we failed to replicate these findings for the same compensation measure, with very weak effect in both experiments (Exp1: $d = 0.09$ [−0.14, 0.31]; Exp2: $d = 0.05$ [−0.16, 0.28]). However, there was a stronger effect for the added regret measure, which was more in line with the original theory (Exp1: $d = 0.50$ [0.27, 0.72]; Exp2: $d = 0.56$ [0.33, 0.79]).

Discussion

Direct replications

The main goal of this study was to replicate and extend three classic experiments that demonstrated stronger regret for exceptions than for routines. In summary, in all three replication studies we found support for an exceptionality effect. However, not all original findings replicated successfully. Scenarios from Part 1 and 2 (hitchhiker- and car accident-scenario) did replicate

successfully, however, the study from Part 3 (robbery scenario) failed to replicate in its original setup and only showed an effect using our added regret measure. A summary of all results is provided in Table 4.

Effect sizes

The patterns of response found in Part 1 and 2 were very similar to the original findings in both replications. In the original hitchhiker-scenario, 88% of the participants rated stronger regret for exceptional behaviour ($d = 2.29$). In the present replication study, 92% of all subjects indicated stronger regret for the exception ($d = 3.12$). In the original car accident-scenario, 82% rated the accident on the unusual route as more regretful ($d = 1.62$), whereas the replication revealed 81% in favour of the exception option ($d = 1.58$).

The scenario used for Part 3 of the present study (robbery-scenario by Miller & McFarland, 1986) was not replicated when using the original study setup. In the original experiment, participants assigned higher amounts of compensation to a victim of a gunshot when the incident happened under

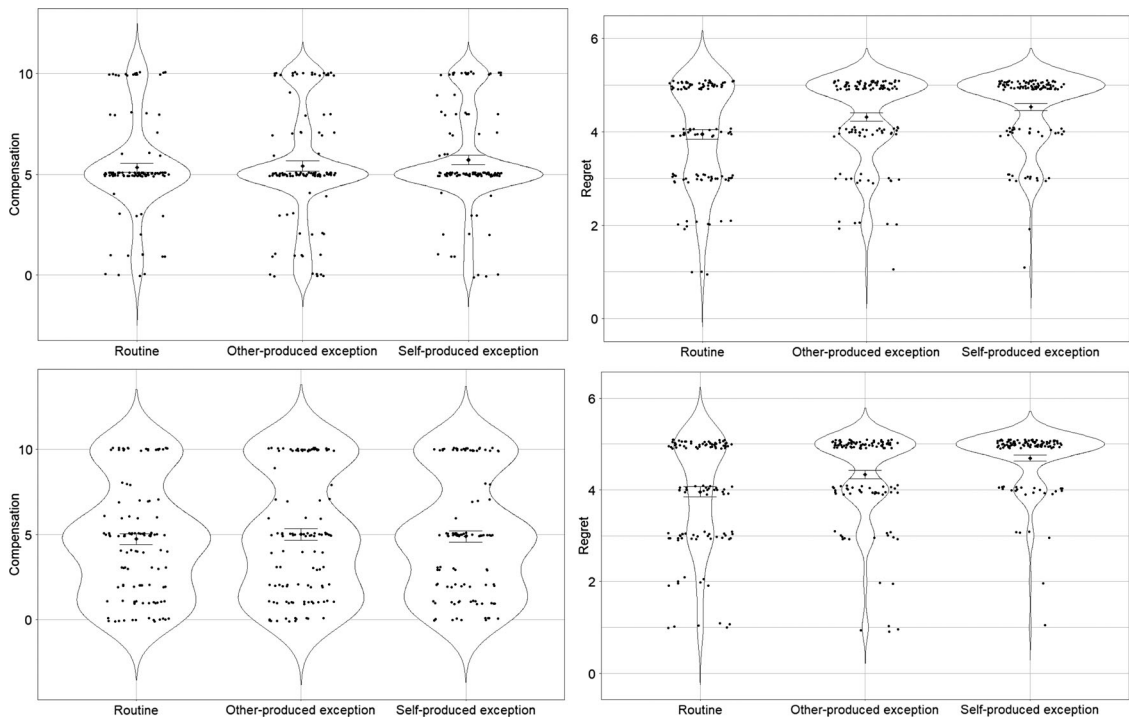


Figure 4. Part 3: Victim compensation and regret by conditions. The top two plots are for Experiment 1, the bottom two plots are for Experiment 2. Error bars represent standard error. Scales: Regret – 1 to 5; Compensation – 1 to 10.

exceptional circumstances ($d = 0.36$, considered a weak to medium effect). Our replication of this experiment revealed no differences in compensation to a victim between routine and exceptional conditions ($d = 0.09$ and 0.05 , very weak effects). However, the regret measure that was added as an extension to the direct replication demonstrated the expected exceptionality effect. Participants assigned stronger

regret to a person when the incident happened in an unusual store in comparison to the victim’s usual market ($d = 0.50$, a medium effect).

Designs

Parts 1 and 2 showed much larger effect sizes than Part 3 with very weak to medium effects. One possible reason for this variance is the difference in design.

Table 4. Summary of experiments and main findings.

Scenario	DV Measure	Routine	Exception	Effect size d	Effect size original	Contribution
Hitchhiker	Regret	7.9%	92.1%	3.12	2.29	Replication
	Social norms 1	4.7%	95.3%	4.29		Extension
	Social norms 2	90.6%	9.4%	2.79		Extension
	Negative affect	7.3%	92.7%	3.28		Extension
Car accident	Regret	19.0%	81.0%	1.58	1.62	Replication
	Randomness	5.67 (1.23)	5.74 (1.23)	0.06		Extension
	Luck	33.3%	66.7%	0.71		Extension
Robbery	Exp1-Compensation	5.34 (2.37)	5.57 (2.61)	0.09	0.36	Replication
	Exp1-Regret	3.95 (1.12)	4.43 (0.87)	0.50		Extension
	Exp2-Compensation	4.74 (3.40)	4.95 (3.57)	0.05	0.36	Replication
	Exp2-Regret	3.95 (1.19)	4.50 (0.86)	0.56		Extension
				[−0.14, 0.31]	[0.03, 0.68]	
				[0.27, 0.72]		

Note: Values in parentheses indicate standard deviation. Values in brackets are 95% confidence intervals (CIs). Effect size d = in comparison measures, converted from chi-square (therefore, with no CIs). Scales: Part 2 Randomness – 1 to 7; Part 3 Compensation – 1 to 10; Part 3 Regret – 1 to 5; All other measures were comparisons.

Both the hitchhiker and car accident scenarios presented two different types of behaviour simultaneously and contrasted the fate of both protagonists. The strong contrast and the direct comparison between exception and routine may have triggered a response in the expected direction. This design helps to illustrate the bias but may suffer from low external validity by using artificial comparison scenarios that are less realistic. In everyday life, people rarely have the opportunity to directly compare the fates of two different persons with the same circumstances and unfortunate outcomes. The robbery scenario uses a between-subject design, in which participants evaluated only a single incident either involving routine or exception. Several studies that tested the exceptionality effect with comparison scenarios demonstrated the bias with very large effects (Briazu et al., 2017; Epstein et al., 1992; Gutten-tag & Ferrell, 2004; Landman, 1987), yet studies that used experiments with between-person designs mostly found medium effects (Macrae, 1992; Macrae et al., 1993; Turley et al., 1995). A study by N'gbala and Branscombe (1997) tested the effect of action-inaction asymmetries on regret (also referred to as the *action-effect*). They found an action-effect when the fate of two persons was presented together, finding stronger regret for action than for inaction. Yet, when using a between-person design, the effect was very weak, demonstrating that scenarios with simple comparisons might over-estimate the effect. Consequently, the magnitude of the exceptionality effect may be much weaker than can be expected from Scenarios 1 and 2 (hitchhiker and car accident).

Failure to replicate Part 3

We failed to replicate the findings in the robbery scenario using the original compensation measure. There could be a number of reasons for that. The replication failure of Part 3 ($d = 0.05$ to 0.09) might indicate that the effect found in the original study was an over-estimation ($p = .03$; $d = 0.36$). Studies on the action-effect and the omission-bias regarding action-inaction asymmetries in the experience of blame and regret highlighted differences between measures of regret and assignment of blame, which is why we added the measure of regret to complement the measure of compensation. We did find the expected effect for the regret measure, and it is reasonable that the exceptionality effect would be more closely associated with regret and counterfactual thought than the

assignment of blame. Regret is an emotion that occurs when people realise that things could have been better if they had decided differently. A measure of regret might, therefore, focus on the emotional response of the victim. Compensation, however, differs from regret in crucial aspects. When asked to assign victim compensation, people might be motivated to assess the caused harm as objectively as possible, anticipating the role of a judge or jury in a process and taking contextual factors into account. The original question also makes it unclear who is to pay the compensation. A study by Turley et al. (1995) demonstrated a related result. Participants rated that a victim of a rape incident would perceive stronger regret when it happened on an unusual route home compared to the same incident on the usual route. The subjects also assigned the victim higher amounts of compensation, but the effect was stronger for regret. Either way, the type of chosen dependent measure might have a crucial moderating role in explaining the magnitude of the exceptionality effect.

Another difference between the studies lies in the reproducibility of Part 1 and 2 (hitchhiker and car accident) in comparison to Part 3 (robbery). The original studies of Part 1 and 2 provided all necessary materials that were needed to conduct a direct replication, including the originally used scenarios. In the case of Part 3, the authors of the original study did not share the original scenario in their publication. Therefore, we reconstructed the scenario as closely as possible using the descriptions from the methods section. Although we are confident that the scenario was rebuilt with very high accuracy taking into account all important facets of the story, details in the vignette description can still impact the study results in an undesired way.

Extensions and moderators

Norm type

In addition to the direct replications, we tested several variables as extensions. In the hitchhiker-scenario used in Part 1, we assessed social normality. We asked the participants who of the two men would more likely be criticised for their behaviour (injunctive social norms) and whose behaviour would be considered more normal in society (descriptive social norms). In support for norm theory, studies have found an impact of social normality on perceptions of regret (Feldman & Albarracín, 2017; Koehler &

Prentice, 2003) finding more regret for socially abnormal behaviour.

In this study, social normality had no direct impact on regret perception. An overwhelming majority of people rated taking hitchhikers as more socially abnormal (injunctive norms: 95%, descriptive norms: 91%²) but people associated stronger regret with the person who took hitchhikers for the first time. In this specific scenario, the deviation from past behaviour by Mr. Jones was more salient in the scenario leading to a stronger impact of past behaviour norms on regret. When two normality types interfere, one norm can outweigh the other.

Are past behaviour norms stronger than social norms? Contrary to our results, research by Hur et al. (2009) contrasted the impact of two normality types and found a stronger influence of social norms on regret when compared to past behaviour norms. Feldman (2018) found comparable effects for past behaviour normality and social norms normality. It is possible that the impact of these two types of norms varies depending on a combination of the routine-exception behaviour, the strength of the social norms, and how negative the outcome is.

Routine strength

The reason for the stronger impact of past behaviour normality on regret in the hitchhiker-scenario might not stem from the type of norm but from the strength of the norm routine. Past behaviour norm, for example, can be very strong when it is performed as a strict routine, such as always driving home the same route. A weaker form of past behaviour would be a usual behaviour, for example, if someone regularly, but not always, goes shopping at the same store. Past behaviour norm can also evolve from a simply repeated behaviour. It can be represented by actions that have been repeated several times but are not performed in a usual manner, such as visiting a certain restaurant occasionally. Following the norm theory logic, an exception might be more abnormal for someone who performs a very strict routine in comparison to somebody who only has repeated a certain behaviour several times.

A similar principle might also apply to social norms normality. Behaviours that are unaccepted in society might lead to strong norms, whereas other norms are weaker or may only play a role in certain parts of society. Regularly taking hitchhikers might not be the most common behaviour but still relatively acceptable in comparison to, for example, criminal activities

such as a robbery or beating up someone in a generally peaceful society.

In the hitchhiker-scenario, past behaviour norm of Mr. Smith who “almost never takes hitchhikers” might be stronger than the social norm that claims that taking hitchhikers as unusual. Therefore, the deviation from the past behaviour might be more salient and has stronger impact on perceived regret.

Randomness and luck

The replication in Part 2 (car accident) was extended with additional measures of random chance and luck in reference to a study by Miller et al. (1989). They found that people evaluated an event as more normal the more similar instances they were able to imagine. Following this logic, people may see an accident on a usual route as more normal than an accident on an unusual route because it is driven more often and, therefore, reveals more paths that could lead to an accident (under the assumption of the same accident probability for both routes).

In the present study, participants evaluated an accident on a routine route as just as likely to occur as an accident on an exceptional route. This result also finds support in studies of Turley et al. (1995) and Macrae (1992) in which participants rated the random chance of a negative incident similar for exceptional and routine routes. People anticipated that the likelihood of a negative incident does not depend on the past behaviour of someone. Nevertheless, most people (81%) rated Mr. White’s exceptional behaviour as more regretful. Although people are capable of estimating the probability of the given event correctly, their judgments of normality are based on the actual number of paths that can lead to a certain outcome.

However, when participants were asked “Which of the two do you think is less lucky?”, the majority of people rated Mr. Jones, who drove the exceptional route, as more unlucky. The difference between the two measures might lie in the measuring method of the two variables. The random chance measure was assessed with two 5-point scales (one for each protagonist) so that people were able to assign similar values to both participants. The luck measure was dichotomous and forced people to decide for one of the two options (Mr. Smith more unlucky vs. Mr. Jones more unlucky). In doubt, people probably again anticipated the protagonist’s emotional reaction and decided that Mr. Jones (exception) must be the less lucky one.

Summary and implications for future research

We conducted three direct replications of classic experiments of the exceptionality effect. The hitchhiker and car accident scenarios (Kahneman & Miller, 1986) were fully replicated. The robbery scenario (Miller & McFarland, 1986) failed to replicate using the original experimental paradigm using a compensation measure, yet we found support for an exceptionality effect using an added measure of perceived regret.

Scenarios of Part 1 (hitchhiker) and 2 (car accident) used comparison scenarios that result in large effect sizes, and the scenario of Part 3 (robbery) showed medium effects. We argued that the medium effect may be closer in magnitude to the “real” effect because comparison measures may inflate the effect sizes of the bias.

We suggested that the exceptionality effect is sensitive to dependent measure (compensation versus regret) and design (comparison versus between-subjects design). Future meta-analysis studies with aggregated power may serve as a tool to clarify the “true” magnitude of the exceptionality effect and test the moderating effect of different designs and measures.

A limitation of this study was that we had no access to the original scenario for Part 3 (robbery). The story was not provided in the original research report and needed to be reconstructed as closely as possible. This emphasises the urgent need for higher transparency and shared documentation in psychological research, to make future work as reproducible as possible.

The analysis of a social normality measure in hitchhiker-scenario showed that the exceptionality effect persisted even when the exceptional behaviour was normal in society. This result may seem contradictory to findings that showed stronger impact of social norms on regret in comparison to past behaviour (Feldman, 2018; Hur et al., 2009). We argued that not only the type but mainly the strength of a norm might have a considerable impact on normality asymmetries on regret. A deviation from a strict daily routine (e.g. always driving home the same route since several years) should be perceived as more abnormal than a behaviour that is inconsistent with behaviour that was sometimes repeated in the past (e.g. visiting a certain restaurant occasionally). The normality strength perspective has not received much attention in the research on norm theory and should be a subject of future work.

Notes

1. Google Scholar search of “(exception OR exceptional OR routine OR normal OR normality) AND (regret OR counterfactual OR counterfactuals)” in papers citing Kahneman and Miller (1986) and Miller and McFarland’s (1986).
2. In the original study, 77% of the participants rated the behaviour of Mr. Smith as more socially abnormal. The difference between replication and original results might lie in the nature of social norms. Social norms are highly culture and context dependent. Canadian participants from the original study in 1986 might, therefore, have a slightly different opinion about taking hitchhikers than participants from the USA in the year 2017 (replication).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gilad Feldman  <http://orcid.org/0000-0003-2812-6599>

References

- Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition*, *167*, 25–37.
- Briazu, R. A., Walsh, C. R., Deeprouse, C., & Ganis, G. (2017). Undoing the past in order to lie in the present: Counterfactual thinking and deceptive communication. *Cognition*, *161*, 66–73.
- Buck, M. L., & Miller, D. T. (1994). Reactions to incongruous negative life events. *Social Justice Research*, *7*, 29–46.
- Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*, 112–130.
- Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, *12*, 105–109.
- 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.
- Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*, 168–192.
- Feldman, G. (2018). What is normal? Contrasting normality categories in the action-effect. Manuscript under review. Retrieved from https://osf.io/wmkpe/?view_only=4d181e034b974b59890650ed0f45d58b
- 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.
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314–325.
- Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, 40, 764–775.
- 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.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136–153.
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–210). Cambridge, MA: Cambridge University Press.
- 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. (2018). The impact of past behavior normality on regret: A meta-analysis. Manuscript under review. Retrieved from https://osf.io/xndtc/?view_only=1c0ed876e0604a5383d013b4dfd3f0dc
- 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.
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49, 433–442.
- 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., & Griffiths, R. J. (1993). Counterfactual thinking and the perception of criminal behaviour. *British Journal of Psychology*, 84, 221–226.
- McEleney, A., & Byrne, R. M. (2006). Spontaneous counterfactual thoughts and causal explanations. *Thinking & Reasoning*, 12, 235–255.
- 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., Taylor, B., & Buck, M. L. (1991). Gender gaps: Who needs to be explained? *Journal of Personality and Social Psychology*, 61, 5–12.
- Miller, D. T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. *Journal of Personality and Social Psychology*, 57, 581–589.
- 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.
- 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.
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133–148.
- Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. *Advances in Experimental Social Psychology*, 56, 1–79.
- Roese, N. J., & Hur, T. (1997). Affective determinants of counterfactual thinking. *Social Cognition*, 15, 274–290.
- 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.
- Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. *Advances in Experimental Social Psychology*, 29, 1–59.
- Turley, K. J., Sanna, L. J., & Reiter, R. L. (1995). Counterfactual thinking and perceptions of rape. *Basic and Applied Social Psychology*, 17, 285–303.
- Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology*, 53, 421–430.

Supplementary Materials

Table of Contents

Power Analyses	2
Open Science.....	2
Procedure and data disclosures	2
Data collection.....	2
Conditions reporting.....	2
Data exclusions.....	2
Variables reporting	3
Compensation Experiments 1 and 2.....	3
Pre-registration Experiment 1.....	4
Hypotheses.....	4
Methods.....	5
Analysis plan.....	6
Additional hypotheses and analyses beyond replication	6
Materials used in Experiment 1.....	7
Scenario #1.....	7
Scenario #2.....	8
Scenario #3.....	8
Pre-registration Experiment 2.....	11
Hypotheses.....	11
Methods.....	13
Analysis plan.....	14
Additional hypotheses and analyses beyond replication	14
Materials used in Experiment 2.....	15

Power Analyses

The effect sizes of the three original experiments were determined. In Part 1 (hitchhiker-scenario) a total of 88% of 138 participants regretted a negative outcome more when the behavior of the focal person was exceptional, resulting in a chi-square effect of 78.38 (compared to a 50-50 random distribution). Chi-square was converted into a Cohen's d of 2.29. In Part 2 (car accident-scenario) a total of 82% of 92 participants regretted a negative outcome more when the antecedent was exceptional resulting in a chi-square effect of 36.57 (compared to a 50-50 random distribution). Chi-square was converted into a Cohen's d of 1.58.

Part 3 originally used three conditions. The first condition was framed as a normal condition (Condition 1: $N=58$, $M=4.52$), the second as self-produced exception condition (Condition 2: $N=48$, $M=5.37$), and the third as other-produced exception condition (Condition 3: $N=57$, $M=5.37$). A preliminary analysis revealed no significant differences ($t < 1$) between the two abnormal conditions. Therefore these conditions were collapsed for the analysis in the original article. T-test result of normal vs. abnormal conditions revealed that the compensation assigned to the victim in the abnormal condition was significantly higher than in the normal condition: ($t(162)=2.17$, $p<.03$). The t-test results were calculated into Cohen's d of 0.36 CIs [0.03; 0.68]

1 - Calculator: <http://www.uv.es/~friasnav/TEdatospublicados.xls>, How to Calculate Effect Sizes from Published Research: A Simplified Spreadsheet By Will Thalheimer and Samantha Cook

2 - MBESS R package. Using $t(162)=2.17$, $p<.03$ and $n_1=58$, $n_2=105$:
library(MBESS); ci.smd(ncp=2.17, n.1=58, n.2=105, conf.level=0.95):
\$Lower.Conf.Limit.smd [1] 0.03148629 \$smd [1] 0.3550135 \$Upper.Conf.Limit.smd [1] 0.6774542

Experiment 3 revealed the smallest effect size ($d = 0.36$) of the three replications and therefore served as the basis for the required sample size. In order to achieve a power of .95 with an alpha of .05 we calculated a minimum need of 336 participants using G* Power Version 3.1.9.2. Our experiments were finally conducted with a total of 342 participants.

Open Science

Data and code will be shared using the Open Science Framework. Files are available for review using the following links:

Experiment 1: <https://osf.io/za7t2/> DOI: 10.17605/OSF.IO/ZA7T2

Experiment 2: - <https://osf.io/p4rgd/> DOI 10.17605/OSF.IO/P4RGD

Procedure and data disclosures

Data collection

Data collection was completed before conducting an analysis of the data.

Conditions reporting

All collected conditions are reported.

Data exclusions

There were no data exclusions. All data is included in the provided data.

Variables reporting

All variables collected for this study are reported and included in the provided data.

Compensation Experiments 1 and 2

Participants received 0.35USD (for a ~2-3 minutes survey).

We included a question in the demographics section of the survey regarding satisfaction with pay for the MTurk HIT, and mean satisfaction was 4 and 4.7 (SD = 1.52 and 1.31) on a 0-6 scale (0 = Extremely unsatisfied, 3 = Neutral, 6 = Very satisfied), with 84% and 95% rating 3-6 (Experiment 1 / Experiment 2, respectively).

Pre-registration Experiment 1

We pre-registered the experiment on May 7th 2017 on the Open Science Framework and data collection was launched later that day. The section pertaining to the current manuscript included the following below:

Hypotheses

Description of essential elements

Norm theory by Kahneman and Miller (1986) proposed that abnormal behavior makes it relatively easy to think of what might have been. In fact, people will perceive higher regret on an episode with a negative outcome when they derive from an intrapersonal norm. The aim of this study is to replicate the effects of three classic experiments:

- Experiment 1 (hitch-hiker scenario) and Experiment 2 (car accident scenario) is a replication of scenarios and data presented in the Kahneman and Miller (1986) review paper.
- Experiment 3 (supermarket scenario) is a replication of Study 1 in Miller & McFarland (1986).

Main Hypothesis:

We expect an impact of intrapersonal norms on perceived regret over a negative outcome. In specific, it is assumed that people would evaluate a person deviating from his or her own past behavior norm as experiencing higher regret in comparison to a person who acted according to his or her own past behavior norm.

Methods

Design

- Part 1
 - Single scenario IV: normal versus abnormal (compared to past behavior). So there's no manipulation, but rather a comparison to a 50-50 split for a random choice.
 - DV: regret
- Part 2
 - Single scenario IV: normal versus abnormal (compared to past behavior). So there's no manipulation, but rather a comparison to a 50-50 split for a random choice.
 - DV: upset
- Part 3
 - IV: 3 conditions between-subject, normal versus abnormal-intentional versus abnormal-unintentional (compared to past behavior)
 - DV:
 - Compensation
 - Regret

Planned Sample

- All experiments will be run with 336 participants from the USA recruited online by using Amazon Mechanical Turk. The sample size was determined through a power analysis based on the effect sizes found in the classic experiments (Power: $1-\beta = 0.95$, Significance: $\alpha = 0.05$). The complete power analysis is provided in Appendix 1. The survey will be pretested with 10 participants on Amazon Mechanical Turk (just to ensure no technical problems, no data peeking).

Exclusion Criteria

We will determine exclusions based on:

- All participants indicating a low proficiency of English (self-report < 5)
- Participants who self-report not being serious about filling in the survey (self-report < 5).

In any case, we will report exclusions in detail with results for full sample and results following exclusions (in either the manuscript or the supplementary).

Procedure

A Qualtrics survey will be used for this study. The survey design is attached to the project to reconstruct the idea. See attached exported Qualtrics survey for full procedure and materials.

Analysis plan

1. Experiment 1 and 2 will be analyzed by using a Chi-square test. It will be assessed if the distribution significantly deviates from a distribution with random chance ($p=.5$).
2. Experiment 3:
 - a. One-way ANOVA of the three conditions with t-test contrasts on all DVs.
 - b. Replicated study methods: If no differences between condition 2 and 3 (conditions for abnormal), these will be combined and a t-test comparison will be performed between the abnormal (2 and 3) versus normal (1).

Additional hypotheses and analyses beyond replication

Going beyond the replication we added new measures on separate pages after the replication:

1. Experiment 1:
 - a. Social norms. Based on Kahneman and Miller (1986) we expect the person regularly taking hitch-hikers will be less common and more criticized by society. Chi-square analyses.
 - b. Negative affect considering both social norms and past behavior: Exploratory, no specific hypothesis. Two-way chi-square analyses (past behavior + perceived norms -> negative affect).
2. Experiment 2, added Miller, Turnbull, and McFarland (1989) related measures about chance and luck.
 - a. Coincidence: Likelihood that accident is random chance for each of the described persons. Within t-test comparison.
 - b. Luck: Which of the two is less lucky: abnormal will be considered less lucky? Chi-square.

Materials used in Experiment 1

Scenario #1

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.

Comprehension questions

Who almost never takes hitch-hikers in his car?

Mr. Jones (1)

Mr. Smith (2)

Who frequently takes hitch-hikers in his car?

Mr. Jones (1)

Mr. Smith (2)

Who got robbed?

Mr. Smith (1)

Mr. Jones (2)

Both Mr. Smith and Mr. Jones (4)

Regret

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

Mr. Jones (1)

Mr. Smith (2)

Social Norms 1

Whose behavior do you think is more common in society?

Mr. Jones (1)

Mr. Smith (2)

Social Norms 2

Whose behavior do you think will be more criticized by others in society?

Mr. Jones (1)

Mr. Smith (2)

Negative affect

Contemplating your previous answers about this scenario and factoring in both Mr. Jones and Mr. Smith personal routines and your perceptions of social norms and possible social criticism, who do you think overall experienced more negative feelings about the decision to take a hitch-hiker that day?

Mr. Jones (1)

Mr. Smith (2)

Scenario #2

Mr. Adams was involved in an accident when driving home after work on his regular route. Mr. White was involved in a similar accident when driving on a route that he only takes when he wants a change of scenery.

Comprehension questions:

Who was driving home after work on his regular route?

Mr. Adams (1)

Mr. White (2)

Who was driving on a route that he only takes when he wants a change of scenery?

Mr. Adams (1)

Mr. White (2)

Who was involved in an accident?

Mr. Adams (1)

Mr. White (2)

Both Mr. Adams and Mr. White (4)

Regret:

Who is more upset over the accident?

Mr. Adams (1)

Mr. White (2)

Random chance:

Please rate your agreement with the following statements:

Mr. Adam's accident is just a random coincidence (1 – strongly disagree to 7 – strongly agree)

Mr. White's accident is just a random coincidence (1 – strongly disagree to 7 – strongly agree)

Luck:

Which of the two do you think is less lucky?

Mr. Adams (1)

Mr. White (2)

Scenario #3

Condition 1

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store A. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Condition 2

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store B because he wanted a change of pace. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Condition 3

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store B because Store A was temporarily closed for renovations. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Comprehension questions:

Which convenience store does Mr. Paul visit frequently?

Store A (1)

Store B (2)

Which convenience store did Mr. Paul visit last night?

Store A (1)

Store B (2)

Did Mr. Paul lose the use of his right arm as a result of a gunshot wound?

Yes (1)

No (2)

How much money should Mr. Paul receive in compensation for his loss?

0 \$ (1)

100k \$ (2)

200k \$ (3)

300k \$ (4)

400k \$ (5)

500k \$ (typical award) (6)

600k \$ (7)

700k \$ (8)

800k \$ (9)

900k \$ (10)

1,000k \$ (11)

Assume there was no compensation given to Mr. Paul. How much regret does he feel over the situation?

no regret (0)

weak regret (1)

medium regret (2)

strong regret (3)

very strong regret (4)

Pre-registration Experiment 2

We pre-registered the experiment on 2018-02-19 12:59 UTC on the Open Science Framework and data collection was launched later that day. The section pertaining to the current manuscript included the following below:

Hypotheses

Description of essential elements

Norm theory by Kahneman and Miller (1986) proposed that abnormal behavior makes it relatively easy to think of what might have been. In fact, people will perceive higher regret on an episode with a negative outcome when they derive from an intrapersonal norm.

In the first experiment we ran the following three scenarios:

- Experiment 1 (hitch-hiker scenario) and Experiment 2 (car accident scenario) are a replication of scenarios and data presented in the Kahneman and Miller (1986) review paper.
- Experiment 3 (supermarket scenario) is a replication of Study 1 in Miller & McFarland (1986).

We successfully replicated the first two, but failed to show a direct replication of Study 1 in Miller & McFarland (1986) using their compensation DV.

The purpose of this study is to again attempt to replicate Study 1 in Miller & McFarland (1986). Since we already conducted a pre-registration for the first attempt to replicate, we note the following changes from the pre-registration of the first attempt:

1. This experiment will focus solely on Study 1 in Miller & McFarland (1986) with no additional studies bundled together, to address possible order effects or that the other experiments somehow affected the experiment.
2. We removed the comprehension checks prior to answering the DV, to address possible effects that these questions may have had on answering the dependent variable.
3. Although the original article indicated a typical compensation award of 500,000 we decided to remove the reference to the typical award from the scale description, as in the previous replication attempt, most participants seemed to simply result to these typical ones. Also, in the previous version we shortened 500,000 to 500k, but in this replication attempt we display the full numerical values, to more closely resemble the original study.

Also, although we failed to replicate the effect, we did find support for the manipulation affecting a regret measure DV which asked "assume there was no compensation given to Mr. Paul. How much regret does he feel over the situation?" (1 – no regret to 5 – very strong regret). In this experiment we made slight changes in the framing of the question to ask about the decision, rather than the situation " Assume there was no compensation given to Mr. Paul. How much regret does he feel about visiting store A?". Also, related to a different project we have on free will attributions, we expected that exceptionality (deviation from routine) would be attributed higher free will than routine, and therefore added an additional DV.

Hypotheses

From the previous pre-registration: We expect an impact of intrapersonal norms on perceived regret over a negative outcome. In specific, it is assumed that people would evaluate a person deviating from his or her own past behavior norm as experiencing higher regret in comparison to a person who acted according to his or her own past behavior norm. New in this study: self-induced exceptionality (a chosen deviation from routine) would be attributed higher free will than routine. In terms of types of exceptionality, we expect that self-produced exception would be the highest free will, compared to other-produced exception. We entertain competing hypotheses regarding the comparison between routine and other-produced exception (exploratory).

Methods

Design

- Experiment
 - IV: 3 conditions between-subject, normal versus abnormal-intentional versus abnormal-unintentional (compared to past behavior)
 - DV:
 - Compensation
 - Regret
 - Free will attributions. This measure of free will attribution (adapted from Clark et al., 2014) will be assessed with three items rated on a 7-point scale.

Planned Sample

- All experiments will be run with 336 participants from the USA recruited online by using Amazon Mechanical Turk. The sample size was determined through a power analysis based on the effect sizes found in the classic experiments (Power: $1-\beta = 0.95$, Significance: $\alpha = 0.05$). The complete power analysis is provided in Appendix 1. The survey will be pretested with 10 participants on Amazon Mechanical Turk (just to ensure no technical problems, no data peeking).

Exclusion Criteria

We will determine exclusions based on:

- All participants indicating a low proficiency of English (self-report < 5)
- Participants who self-report not being serious about filling in the survey (self-report < 5).

In any case, we will report exclusions in detail with results for full sample and results following exclusions (in either the manuscript or the supplementary).

Procedure

A Qualtrics survey will be used for this study. The survey design is attached to the project to reconstruct the idea. See attached exported Qualtrics survey for full procedure and materials.

Analysis plan

Experiment 3:

1. One-way ANOVA of the three conditions with t-test contrasts on all DVs.
2. Replicated study methods: If no differences between condition 2 and 3 (conditions for abnormal), these will be combined and a t-test comparison will be performed between the abnormal (2 and 3) versus normal (1).
3. Free will attributions: We will average the 3 items, if the Cronbach's alpha is superior or equal to .70. If the Cronbach's alpha is inferior to .70, we will analyze the items independently.

Items (1 = strongly disagree, 7 = strongly disagree):

- a. Mr. Paul could have chosen to do otherwise and visit a different store
- b. By visiting that specific store, Mr. Paul was exercising his free will
- c. Mr. Paul's choice of which store to visit was his own free choice

Criteria: NHST, using $\alpha < .05$, one-tail. This is to mirror the original studies.

Additional hypotheses and analyses beyond replication

Going beyond the replication we added new measures in separate pages after the replication:

1. In accordance with previous pre-registration: added a measure of regret
2. New: added free will attributions. (to serve a different project)

Materials used in Experiment 2Condition 1

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store A. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Condition 2

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store B because he wanted a change of pace. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Condition 3

Two convenience stores are located in Mr. Paul's neighborhood. He frequents Store A more regularly than Store B. Last night he visited Store B because Store A was temporarily closed for renovations. He walked in on a robbery taking place at the store, and was shot. He lost the use of his right arm as a result of the gunshot wound.

Mr. Paul seeks compensation for both the physical and psychological harm suffered.
How much money should Mr. Paul receive in compensation?

- 0 \$ (1)
- 100k \$ (2)
- 200k \$ (3)
- 300k \$ (4)
- 400k \$ (5)
- 500k \$ (typical award) (6)
- 600k \$ (7)
- 700k \$ (8)
- 800k \$ (9)
- 900k \$ (10)
- 1,000k \$ (11)

Assume there was no compensation given to Mr. Paul. How much regret does he feel about visiting store [A/B]?

- no regret (0)
- weak regret (1)
- medium regret (2)
- strong regret (3)
- very strong regret (4)