Article

Social Psychological and Personality Science
2021, Vol. 12(7) 1146-1154

DOI: 10.1177/1948550620966153
journals.sagepub.com/home/spp
SAGE

Pupillometry and Hindsight Bias: Physiological Arousal Predicts Compensatory Behavior

Willem W. A. Sleegers[1], Travis Proulx[2], and Ilja van Beest[1]

Abstract

According to violation–compensation models of cognitive conflict, experiences that violate expected associations evoke a common, biologically based syndrome of aversive arousal, which in turn motivates compensation efforts to relieve this arousal. However, while substantial research shows that people indeed respond with increased arousal to expectancy violating events, evidence for the motivating role of arousal is rarely found. In two within-subjects studies ($N = 44$ and $N = 50$), we demonstrate evidence for the motivating role of arousal in this violation–compensation process among university students. Using pupillometry and the hindsight bias phenomenon, we show that people respond with greater arousal when presented with expectancy violating information. In turn, we show that the pupillary response is positively related to the amount of hindsight bias being displayed. These findings provide further insights into the process underlying the hindsight bias and, crucially, support key predictions following from threat–compensation models.

Keywords

threat–compensation, arousal–behavior link, pupillometry, hindsight bias

Humans adopt a multitude of beliefs ranging from worldviews to trivial facts. They may believe that the world is fair, that behavior follows from attitudes, and that people eat an average of eight spiders per year in their sleep. This range of beliefs reflects a ubiquitous need for consistent *meaning*, the worldviews that organize our perception of the world (Meaning Maintenance Model; Heine et al., 2006). These worldviews are comprised of expected relations between experiences, and when these expectations are met, a sense of meaning is experienced.

But these expectations are not always correct. It turns out the world is not always fair (Lerner, 1980) and that people do not eat eight spiders in their sleep on an annual basis (Sneed, 2014). These events violate the sense of meaning and cause a state of discomfort. In turn, this discomfort motivates subsequent compensatory behaviors in order to restore meaning (Jonas et al., 2014; McGregor et al., 2012; Proulx et al., 2012). One form of compensatory behavior is to assimilate events so that they appear consistent with initial expectations (Park, 2010; Piaget, 2000). For example, a misfortune that befalls an innocent person can be interpreted as deserving rather than unfair, such as a victim of rape being accused of having provoked it by dressing provocatively. This assimilation maintains a sense of consistency with the belief of a just world (Lerner, 1980). Additional strategies may be used, such as accommodating one's belief to the expectancy violation or by affirming unrelated meaning frameworks (Heine et al., 2006; Proulx & Inzlicht, 2012).

An important tenet of violation–compensation theories is that expectancy violations induce a syndrome of aversive arousal, motivating the execution of compensatory behaviors. If this is indeed the case, then at least two lines of evidence should be found (Townsend et al., 2013). First, expectancy violations should induce a state of heightened physiological arousal. Second, this arousal should be linked to the compensatory behavior. Evidence for the former can be found in abundance, whether it is an expectancy violation caused by perceptual anomalies (Sleegers et al., 2015), cognitive dissonance (Gerard, 1967), self-view inconsistencies (Ayduk et al., 2012), worldview violations (Townsend et al., 2010), or category-based violations (Mendes et al., 2007).

Evidence for the second link is rarely observed. In fact, the evidence for the role of arousal comes mostly from indirect

[1] Department of Social Psychology, Tilburg University, the Netherlands
[2] Cardiff University, United Kingdom

Corresponding Author:
Willem W. A. Sleegers, Department of Social Psychology, Tilburg University, PO Box 90153, 5000 LE Tilburg, the Netherlands.
Email: w.w.a.sleegers@tilburguniversity.edu

assessments of arousal rather than direct measures. Some evidence has been found using self-report measures to index the aversive arousal and its association with compensatory behavior (Laurin et al., 2008; McGregor et al., 2013, Experiment 4; Plaks et al., 2005). For example, Plaks et al. (2005, Experiment 3) found that self-reported anxiety mediated the relationship between expectancy violation and an increased need for certainty.

Others have found indirect support for the second link utilizing the misattribution of arousal paradigm (Kay et al., 2010; Losch & Cacioppo, 1990; Proulx & Heine, 2008; Zanna & Cooper, 1974). For example, Proulx and Heine (2008) presented participants with an implicit perceptual anomaly and administered a placebo. Those who were informed that the placebo caused side effects of arousal did not affirm a valued moral belief compared to those who were not informed of such side effects.

These examples are, however, the exception rather than the rule as evidence for the mediational link of arousal remains elusive (also see McGregor et al., 2013, p. 550). There are multiple potential reasons for why the link between arousal and compensatory behaviors is rarely found. These include methodological reasons such as poor measurement reliability or low power (Mauss et al., 2005) and also reasons directly related to the topic of investigation. For example, self-report measures can provide a compensatory opportunity to respond to the expectancy violation. By explicitly reporting that the violation was not distressing, one can persuade oneself to feel less distress (Elliot & Devine, 1994; Galinsky et al., 2000).

With regard to direct measures of arousal, there may also be a limitation that prevents the discovery of an arousal–compensatory behavior link: Physiological arousal measurement tools themselves evoke arousal. The placing of electrodes on the skin or one's head between metal braces can be arousing or is expected to evoke arousal. Consequently, it is possible that participants attribute arousal caused by a manipulation to the measurement tool, ironically showing the efficacy of misattribution studies that in other contexts serve as evidence for the role of arousal in the threat–compensation process. To illustrate, Croyle and Cooper (1983) performed a standard cognitive dissonance paradigm and found the predicted pattern of attitude change. This effect disappeared in a subsequent study, which contained a physiological arousal assessment in the form of skin conductance. The authors interpreted this absence of an effect due to participants misattributing their arousal to the physiological recording device.

In summary, according to violation–compensation theories of cognitive conflict, expectancy violations induce a state of aversive arousal (Link 1) that motivates compensatory behavior (Link 2). Myriad findings are in support of the first link, but the second link remains uncertain in terms of direct empirical support. In the present article, we investigate this link using a design that takes into account the limitations of prior work. We use eye tracker technology to assess arousal in response to expectancy violations and use repeated measurements of compensatory behavior.

Pupillometry and Psychophysiological Arousal

Modern eye trackers are easier to use and, more importantly, less invasive than before. Unlike previous models that often required the participant's head to be fixed in place, modern eye trackers use screen-based solutions that can record eye properties from a distance, without restraining the participant.

Eye trackers can measure a state of arousal using pupillometry. Pupillometry is the measurement of pupil size and its reactivity. The size of the pupil is about 3 mm in standard lighting conditions and can range from 1.5 mm to 9 mm (Wyatt, 1995). Since the late 1800s, it has been shown that pupil size can serve as a proxy for a state of psychophysiological arousal due to tiny fluctuations that cannot be explained by changes in luminance (Beatty & Lucero-Wagoner, 2000). These variations stem from two smooth muscles in the iris: a sphincter and dilator muscle. The dilator muscle is under adrenergic control by the sympathetic nervous system, which, when active, leads to the stimulation of the dilator muscle and a parallel inhibitory effect on the sphincter muscle via a parasympathetic mechanism (Sirois & Brisson, 2014). The relationship between pupil size and arousal stems from its association with the locus coeruleus–norepinephrine (LC–NE) system (Gilzenrat et al., 2010; Murphy et al., 2014). The LC is a subcortical structure and is the principal site for the production and release of NE (Sara, 2009). The LC–NE is involved in various arousal processes including stress responses, memory retrieval, and attention and more generally underlies the regulation of engagement or withdrawal from a task (for a review, see Aston-Jones & Cohen, 2005). As a result, pupil size increases have been observed following the presentation of positively and negatively valenced pictures (Bradley et al., 2008), the experience of pain (Chapman et al., 1999; Ellermeier & Westphal, 1995; Höfle et al., 2008), task error (Brown et al., 1999; Critchley et al., 2005), and perceptual inconsistencies (Preuschoff et al., 2011; Raisig et al., 2010, 2012; Sleegers et al., 2015)—making pupillometry a valid, and nonintrusive, tool to assess arousal.

A consequence of pupillometry is the need for repeated trials in order to obtain a reliable measure of arousal. This means a within-subjects design is preferred in which the link between arousal and compensatory behavior is repeatedly assessed. The majority of studies on compensatory behavior use between-subject designs in which compensatory behavior is measured once, so using common compensatory behavior assessments is not ideal. Therefore, we will instead rely on a different paradigm to assess compensation behaviors, based on a well-known psychological phenomenon: the hindsight bias.

Hindsight Bias as a Compensatory Response

Hindsight bias, or the "knew-it-all-along" effect, is the tendency for individuals with outcome knowledge (hindsight) to claim that they did in fact know the outcome or estimated its occurrence with a higher probability than they would have estimated without the outcome information (foresight). Several processes have been proposed to underlie the hindsight bias (Hawkins & Hastie, 1990; Roese & Vohs, 2012). It has been suggested that the hindsight bias is caused by a relatively automatic and unconscious sensemaking process (Fischhoff, 1975), and research on individual differences in the tendency to demonstrate hindsight bias has revealed that people high in need of predictability and control more frequently show hindsight bias (Musch, 2003). These findings suggest that motivational processes may underlie the hindsight bias, similar to, or perhaps even identical to, the processes underlying compensatory responses following expectancy violations.

Assessing hindsight bias as a compensatory response has the methodological benefit of allowing for a repeated measures design. Using the so-called memory paradigm, participants first answer a series of questions to establish their prior beliefs, followed by a second presentation of these questions together with their correct answers. After the presentation of the correct answer, the participant is prompted to report their original response. If the recalled response is different from the initial response and closer to the correct answer, a hindsight bias has been demonstrated. For our purposes, the advantage is that there is no limit to how many questions can be asked, except for taking into account participant fatigue and the question pool. This makes it a viable design to repeatedly assess compensatory affirmation behaviors.

Present Studies

In two studies, we use pupillometry and the hindsight bias to investigate the role of physiological arousal in the relationship between expectancy violations and compensatory assimilation behavior. We hypothesize that incorrect answers violate the expectations of the participant, resulting in greater pupil dilation. This increased pupil dilation should motivate participants to indicate a different response than initially given, in the direction of the presented correct answer, that is, display a hindsight bias. Such a relationship would provide the first direct support violation–compensation theories, whereby compensation behaviors are shown to be a direct response to aversive arousal in response to an expectancy violation.

Study 1

Method

Participants

Students ($N = 44$, 31 women, $M_{\text{age}} = 20.93$ years) at Tilburg University participated in exchange for course credit or a monetary reward. The majority of participants (34) were undergraduate students in psychology. Sample size was based on prior research using pupillometry (Bradley et al., 2008; Laeng et al., 2011; Partala & Surakka, 2003). No additional data were collected after data analysis had begun.

Design and Procedure

The present study consisted of a full within-subjects design, with a hindsight bias paradigm to present expectancy violations and to measure compensatory behavior. Specifically, we used a memory hindsight bias design (Calvillo, 2013; Pohl, 2007) in which participants answered a series of factual questions, first before seeing the correct answers and again later after seeing the correct answer to each question. Each question was presented individually. After participants indicated what they believed to be the correct answers, the eye tracker was calibrated and participants saw each question again, followed by the presentation of the correct answer to said question. Pupil size was measured during the presentation of the correct answer. Immediately following the correct answer, the participant had to indicate what their original answer was. At the end, participants filled in several demographic questions.

Materials

Almanac questions. The questions for the hindsight bias task were selected from various online sources and books on the topic of misconceptions (e.g., van Maanen, 1994). We selected 80 almanac questions (Online Appendix A) that we believed participants thought they could answer or guess but that varied in terms of whether they would answer correctly. In other words, we selected questions that varied in the extent that the correct answer would surprise them.

Hindsight bias task. The hindsight bias task consisted of two parts, each consisting of 80 trials. In the first part, a trial consisted of a single question, and participants were asked to indicate what they believed to be the correct answer. Questions were presented in random order, without a time limit. In the second part of the hindsight bias task, a trial consisted of a single question, shown for a minimal duration of 3,000 ms, after which the participant could click with the mouse to continue. Hereafter, there was a blank screen for a duration of 1,000, 1,500, or 2,000 ms, followed by a fixation cross (3,000 ms). After the fixation cross, the correct answer to the quest was presented for a duration of 5,000 ms. Hereafter, participants were asked to indicate what their answer was, identical to that in the first part.

Hindsight bias. Hindsight bias was defined as the difference between the second and the first response to each question, with the requirement that the second response was closer to the correct answer than the first (Pohl, 2007). We created a percentage-based solution to reduce the influence of questions with extremely large numeric answers by dividing the amount of hindsight bias by the absolute distance between the correct answer and the participant's initial response (Hell et al., 1988). This means that correct responses are removed and that typical responses should fall within the range of 0–1. If the hindsight bias is greater than 1, it indicates an overcorrection (e.g., if the first response is 30, the correct answer is 40, and the second response is 45). If the response is smaller than 0, it means the second response was further away from the correct answer than their first response. Some of these responses were likely due to typos (e.g., missing a 0 with large numbers) and guesses. If the participant guessed their first answer, it may be more difficult for them to remember it after seeing the correct answer, resulting in a random change between the two responses. Due to the difficulty of interpreting the meaning of these responses in terms of a hindsight bias, we chose to remove these trials for the main analyses. We report additional analyses in the Supplemental Material (Online Appendix C) in which we included the overcorrection response and set responses smaller than 0 to indicate 0 hindsight bias.

Pupillometry. A Tobii T60 eye tracker (Tobii, Stockholm, Sweden) was used to record pupil data. The Tobii T60 is a noninvasive eye tracker that is integrated in a 17″ thin-film transistor monitor, resembling a standard PC monitor. It records at a rate of 60 Hz. Each measurement has a validity indication that ranges from 0 (the system is certain that all data belong to the particular eye) to 4 (gaze data are missing or incorrect). Only recordings with a validity score of 0 or 1 were used. Following guidelines of Kret and Sjak-Shie (2019), pupil sizes of each eye separately were preprocessed by removing dilation speed outliers, observations near gaps, and sparse clusters (see Online Appendix B for more details). Then, missing data were linearly interpolated and smoothened using a 10-Hz low-pass filter. Hereafter, the pupil size was controlled for baseline differences by subtracting the average pupil size during a 500-ms preceding the answer period from the pupil measurements during the answer period (Beatty & Lucero-Wagoner, 2000). For data analysis, the pupil size was averaged across a period of 500–4,000 ms during the presentation of the correct answer.[1] The initial 500 ms was seen as the light reflex period. Trials with more than 25% missing data were excluded.

Data Analysis

The data were prepared and analyzed in R (R Core Team, 2019). Both the data and the scripts are available at the DataverseNL repository.[2] Below, we report all data exclusions prior to the main analysis. In addition, we test whether pupil dilation was associated with being mistaken in order to validate the response feedback as indeed affecting the participant's physiology. To this end, we created a binary variable that indicated whether the participant gave a correct or incorrect answer on each trial. A binary measure of error was created because the magnitude of the error was heavily dependent on the question that was asked.

Multilevel models were used to test the hypotheses. In each model, we include by-participant and by-question random intercepts. The *lme4* package (Bates et al., 2015) was used for the mixed-model analyses in combination with the *lmerTest* package (Kuznetsova et al., 2016) in order to obtain p values. The normality assumption of the residuals was checked and found to be untenable for the hindsight bias analysis, leading us to also perform a robust multilevel model using the *robustlmm* package (Koller, 2016). Because the *robustlmm* package does not provide p values, we extracted the Satterthwaite approximated degrees of freedom from the nonrobust model and the t value of the robust model to calculate p values.

Results

Data Exclusions

Participants completed a total of 3,520 trials. To assure reliable pupil observations, trials with more than 25% missing data were removed (26.79%), leaving 2,535 trials. Of these trials, five responses were removed because they indicated mistakes (e.g., giving a response higher than 100 when a percentage was asked). Participants were correct on 518 (20.47%) trials, meaning a hindsight bias score could not be determined for these trials. Of the possible hindsight bias trials, participants overcorrected on 31 (1.54%) trials and moved away from the correct answer on 173 (8.60%) of the trials. Excluding these trials resulted in 1,808 trials, of which 366 (21.48%) displayed a hindsight bias.

Pupil Dilation and Error

To test whether being mistaken was associated with an increase in pupil size, we conducted a linear mixed model with average pupil dilation as the outcome variable and error (correct/incorrect) as the predictor. This revealed a significant effect of being mistaken, $b = .020$, 95% confidence interval (CI) = [.0049, .036], $SE = .008$, $t(292.06) = 2.58$, $p = .011$. On trials in which participants were incorrect, greater pupil dilation was observed ($M = .014$) than on trials in which they gave the correct answer ($M = -.0060$).

Pupil Dilation and Hindsight Bias

To test our main hypothesis, we conducted a linear mixed model with hindsight bias as the outcome variable and pupil dilation as the predictor. This revealed a significant relationship between hindsight bias and pupil size, $b = .098$, 95% CI [.032, .16], $SE = .033$, $t(1{,}689.28) = 2.93$, $p = .0034$. Participants' pupillary reaction was positively related to the amount of hindsight bias they displayed.

Inspecting the distribution of the residuals revealed that the normality assumption was violated. We therefore reran the same model using a robust linear mixed model and found that the positive relationship between pupil size and hindsight bias was now marginally significant, $t(1{,}689.28) = 1.93$, $p = .053$, calling for a replication of the current work.

Study 2

Study 2 is a replication of the previous work, with several small changes to improve the previous design. We therefore only note the differences between the two studies before reporting the results.

Method

Participants

Fifty students (35 women, $M_{age} = 20.7$ years) at Tilburg University participated in exchange for course credit. Participant recruitment took place in two phases: at the end of the academic year and at the start of the next academic year. No additional data were collected after data analysis had begun. A difference with Study 1 is that both Dutch ($N = 29$) and international students ($N = 21$) participated.

Materials

Almanac questions. Because international students could participate in the present study, several of the almanac questions from Study 1 were removed in favor of less culture-specific questions. In addition, 20 questions were added to a total of 100 almanac questions, compared to the 80 questions in Study 1.

Pupillometry. In addition to the Tobii T60 eye tracker, we used a Tobii Pro Spectrum eye tracker. The Tobii Pro Spectrum is an eye tracker integrated in a 24″ screen and capable of recording at 150 Hz. However, to match the recording rate of the T60, both eye trackers recorded data at 60 Hz.

Data Analysis

Below we report the results of Study 2, as well as the results of both studies taken together. To this end, we combined the two data sets and conducted the same analyses as in both studies, but with the addition of by-study random intercepts.

Results

Data Exclusions

Participants completed a total of 5,000 trials. As in Study 1, trials with more than 25% missing data were removed (28.20%), leaving 3,544 trials. Of these trials, no responses were removed because they indicated mistakes, but one participant was removed for always responding with the same answer. Participants were correct on 1,167 (33.08%) trials, meaning a hindsight bias score could not be determined for these trials. Of the possible hindsight bias trials, participants overcorrected on 44 (1.86%) trials and moved away from the correct answer on 248 (10.50%) of the trials. Excluding these trials resulted in 2,069 trials, of which 514 (24.84%) displayed a hindsight bias.

Pupil Dilation and Error

Similar to the results in Study 1, we found a significant effect of being mistaken on pupil size, $b = .025$, 95% CI [.012, .038], $SE = .0066$, $t(318.64) = 3.72$, $p < .001$. On trials in which participants were incorrect, greater pupil dilation was observed ($M = .021$) than on trials in which they gave the correct answer ($M = -.0040$).

Pupil Dilation and Hindsight Bias

In contrast to Study 1, we found only a marginally significant relationship between hindsight bias and pupil size, $b = .059$, 95% CI [−.005, .12], $SE = .033$, $t(2{,}024.82) = 1.81$, $p = .07$. We did find that the residuals were again not normally distributed. We therefore reran the same model using a robust linear mixed model and did not find a significant positive relationship, $t(2{,}024.82) = 1.34$, $p = .183$.

Results of Studies 1 and 2

Because Studies 1 and 2 were identical in all relevant respects, we combined the data and reran the analyses for greater statistical power. We expanded the model by including by-study random intercepts to account for the fact that the data stem from two separate studies. This analysis again revealed the effect of being mistaken on pupil size, $b = .023$, 95% CI [.013, .033], $SE = .0051$, $t(486.69) = 4.41$, $p < .001$, and also the positive relationship between pupil size and the hindsight bias from Study 1, $b = .078$, 95% CI [.036, .13], $SE = .024$, $t(3{,}806.61) = 3.31$, $p < .001$. More importantly, the robust mixed model now confirmed the relationship between pupil size and hindsight bias, $t(3{,}806.61) = 2.32$, $p = .021$.

Discussion

We aimed to demonstrate the first direct link between physiological arousal and compensatory behavior. While the results of each study separately were not conclusive, the results from both studies combined did provide evidence for this link. Greater pupil dilation in response to an unexpected correct answer was associated with more hindsight bias. That is, participants shifted their second answer more toward the factual question's correct answer, relative to their first answer, when they showed a larger physiological response to the correct answer to the question. This compensatory response following increased arousal is consistent with violation–compensation theories (Jonas et al., 2014; McGregor et al., 2012), specifically with the shared assumption that inconsistencies evoke arousal that causes compensation reactions.

That expectancy violations induce a syndrome of aversive arousal is an important tenet of violation–compensation theories. There is abundant evidence for this first link between expectancy violations and arousal, whether the expectancy violation involves perceptual anomalies (Sleegers et al., 2015), cognitive dissonance (Gerard, 1967), self-view inconsistencies (Ayduk et al., 2012), worldview violations (Townsend et al., 2010), or category-based violations (Mendes et al., 2007). Evidence for the second link, between arousal and the subsequent compensatory behavior, is rarely observed and limited to indirect assessments of arousal such as self-report measures

(Laurin et al., 2008; McGregor et al., 2013, Experiment 4; Plaks et al., 2005) and the misattribution of arousal paradigm (Kay et al., 2010; Losch & Cacioppo, 1990; Proulx & Heine, 2008; Zanna & Cooper, 1974). Our findings provide more direct evidence for the often postulated relationship between arousal and compensatory behaviors following expectancy violations.

Two reasons might explain why we were able to demonstrate a link between arousal and compensatory behavior. First, recent developments in eye tracker technology have made this technology exceptionally noninvasive. Consequently, an eye tracker is less likely to evoke arousal that interferes with the arousal process underlying violation–compensation reactions. Second, we repeatedly presented participants with an expectancy violation and an opportunity to compensate—a requirement for physiological measures to improve reliability.

Limitations and Future Research

In our studies, we relied on pupillometry to assess an aversive state of arousal following negative belief feedback because threat–compensation theories strictly postulate a state of aversive arousal to motivate subsequent compensatory behaviors. However, while pupillometry is a valid measure of physiological arousal, it is not a direct measure of *aversive* arousal (e.g., Bradley et al., 2008). We believe our findings nevertheless plausibly indicate a state of aversive arousal. Studies have shown that negative belief feedback and states of surprise are (at least initially) experienced as aversive (Hajcak & Foti, 2008; Noordewier & Breugelmans, 2013; Noordewier et al., 2016). In addition, alternative explanations such as curiosity-driven responses were ruled out by the data (see Online Appendix C). We therefore believe our findings present a strong contribution to models of threat–compensation.

It should be noted that we relied mostly on epistemic threats rather than more severe existential threats such as those relating to one's identity or freedom. Epistemic threats were chosen in order to be able to repeatedly present participants with threats and compensation opportunities. This would not be feasible when more impactful threats are used because the physiological response would likely carry over between trials and affect the relationship between arousal and compensation. Moreover, the theoretical perspectives that guide this research share the explicit premise that the response to epistemic threats generalize to other types of threats (Heine et al., 2006; Jonas et al., 2014; Proulx & Inzlicht, 2012). In fact, it has been demonstrated that the experience of inconsistency, such as those experienced by our participants, can evoke the same compensation behaviors as existential threats (e.g., nonsense word pairs and identity violations; Randles et al, 2011). Nevertheless, the threat–compensation literature would benefit from more empirical demonstrations of the kind presented here.

Aside from expectancy violations inducing physiological arousal, and physiological arousal motivating compensatory behavior, compensatory behavior should also reduce the physiological arousal. We did not assess this third link. Using the present studies' design, it might be possible to demonstrate the entire causal link by having participants again see the correct answers. We predict that instead of the positive relationship between pupil size and hindsight bias found in the present study, a negative relationship between hindsight bias and pupil size should be found.

Finally, in the present studies, we used the hindsight bias as a way to repeatedly assess compensatory behaviors following belief violations. It may be argued that due to the many trials, participants may not have always remembered their initial answer and that this ultimately shaped their hindsight bias responses. However, research on the hindsight bias largely supports a biased reconstruction view rather than a memory impairment process (Stahlberg & Maass, 1997). Our findings also contribute to the research on the hindsight bias. Several processes have been proposed to explain the hindsight bias (Hawkins & Hastie, 1990), including motivational accounts (Campbell & Tesser, 1983; Fischhoff, 1975; Musch, 2003). Our results are consistent with a motivational interpretation of the hindsight bias, thereby also contributing to research on the hindsight bias phenomenon.

We did employ a memory design to measure hindsight bias. Importantly, this memory-based design, although effective in demonstrating a hindsight bias, might be less effective in evoking a hindsight bias than other designs such as the hypothetical design (Pohl, 2007), in which participants are asked to respond as if they had not been told the correct answer. After all, a memory task is about recalling a previously reported answer; and when the time lag is not substantial, people can with relative ease recall their answer. For this reason, the memory design can be potentially improved in future studies by extending the retention interval between the first and second responses.

Conclusion

We found that the magnitude of hindsight bias was positively related to the size of pupil change in response to seeing the expected and unexpected correct answer to a set of questions. This finding is consistent with violation–compensation theories that postulate a role of aversive psychophysiological arousal in producing compensatory behavior following expectancy violations.

Declaration of Conflicting Interests

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

Funding

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

ORCID iDs

Willem W. A. Sleegers https://orcid.org/0000-0001-9058-3817
Ilja van Beest https://orcid.org/0000-0003-2855-3638

Supplemental Material

The supplemental material is available in the online version of the article.

Notes

1. We also conducted additional analyses in which we used the pupil period *before* seeing the correct answer in order to rule out a curiosity-driven alternative explanation. See Appendix C for these analyses.
2. For review purposes, please see the following anonymized open software framework page (https://osf.io/fb2nk/?view_only=7132532e4cad49c58752d0f6d7f2cfdb).

References

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Ayduk, O., Gyurak, A., Akinola, M., & Mendes, W. B. (2012). Consistency over flattery: Self-verification processes revealed in implicit and behavioral responses to feedback. *Social Psychological and Personality Science*, *4*(5), 538–545. https://doi.org/10.1177/1948550612471827

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 142–162). Cambridge University Press.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x

Brown, G. G., Kindermann, S. S., Siegle, G. J., Granholm, E., Wong, E. C., & Buxton, R. B. (1999). Brain activation and pupil response during covert performance of the Stroop color word task. *Journal of the International Neuropsychological Society*, *5*(4), 308–319. https://doi.org/10.1017/S1355617799544020

Calvillo, D. P. (2013). Rapid recollection of foresight judgments increases hindsight bias in a memory design. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 959–964. https://doi.org/10.1037/a0028579

Campbell, J. D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal of Personality*, *51*(4), 605–620. https://doi.org/10.1111/j.1467-6494.1983.tb00868.x

Chapman, C. R., Oka, S., Bradshaw, D. H., Jacobson, R. C., & Donaldson, G. W. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, *36*, 44–52. https://doi.org/10.1017/S0048577299970373

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., & Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *NeuroImage*, *27*(4), 885–895. https://doi.org/10.1016/j.neuroimage.2005.05.047

Croyle, R. T., & Cooper, J. (1983). Dissonance arousal: Physiological evidence. *Journal of Personality and Social Psychology*, *45*(4), 782–791. https://doi.org/10.1037//0022-3514.45.4.782

Ellermeier, W., & Westphal, W. (1995). Gender differences in pain ratings and pupil reactions to painful pressure stimuli. *Pain*, *61*(3), 435–439. https://doi.org/10.1016/0304-3959(94)00203-Q

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

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(3), 288–299. https://doi.org/10.1037/0096-1523.1.3.288

Galinsky, A. D., Stone, J., & Cooper, J. (2000). The reinstatement of dissonance and psychological discomfort following failed affirmations. *European Journal of Social Psychology*, *30*(1), 123–147. https://doi.org/10.1002/(SICI)1099-0992(200001/02)30:13.0.CO;2-T

Gerard, H. B. H. B. (1967). Choice difficulty, dissonance, and the decision sequence. *Journal of Personality*, *35*(1), 91–108. https://doi.org/10.1111/j.1467-6494.1967.tb01417.x

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 252–269. https://doi.org/10.3758/CABN.10.2.252

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological Science*, *19*(2), 103–108. https://doi.org/10.1111/j.1467-9280.2008.02053.x

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*(3), 311–327. https://doi.org/10.1037/0033-2909.107.3.311

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review*, *10*(2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1

Hell, W., Gigerenzer, G., Gauggel, S., Mall, M., & Müller, M. (1988). Hindsight bias: An interaction of automatic and motivational factors? *Memory & Cognition*, *16*(6), 533–538. https://doi.org/10.3758/BF03197054

Höfle, M., Kenntner-Mabiala, R., Pauli, P., & Alpers, G. W. (2008). You can see pain in the eye: Pupillometry as an index of pain intensity under different luminance conditions. *International Journal of Psychophysiology*, *70*(3), 171–175. https://doi.org/10.1016/j.ijpsycho.2008.06.008

Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., Nash, K., Proulx, T., & Quirin, M. (2014). Threat and defense: From anxiety to approach. *Advances in Experimental Social Psychology*, *49*, 219–286. https://doi.org/10.1016/B978-0-12-800052-6.00004-4

Kay, A. C., Moscovitch, D. A., & Laurin, K. (2010). Randomness, attributions of arousal, and belief in God. *Psychological Science*, *21*(2), 216–218. https://doi.org/10.1177/0956797609357750

Koller, M. (2016). Robustlmm: An R package for robust estimation of linear mixed-effects models. *Journal of Statistical Software*, *75*(6). https://doi.org/10.18637/jss.v075.i06

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods*, *51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models.* Journal of Statistical Software, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing*, *12*(1), 13–21. https://doi.org/10.1007/s10339-010-0370-z

Laurin, K., Kay, A. C., & Moscovitch, D. A. (2008). On the belief in god: Towards an understanding of the emotional substrates of compensatory control. *Journal of Experimental Social Psychology*, *44*(6), 1559–1562. https://doi.org/10.1016/j.jesp.2008.07.007

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. Springer. https://doi.org/10.1007/978-1-4899-0448-5

Losch, M. E., & Cacioppo, J. T. (1990). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology*, *26*(4). https://doi.org/10.1016/0022-1031(90)90040-S

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*(2), 175–190. https://doi.org/10.1037/1528-3542.5.2.175

McGregor, I., Prentice, M., & Nash, K. (2012). Approaching relief: Compensatory ideals relieve threat-induced anxiety. *Social Cognition*, *30*(6), 689–714. https://doi.org/10.1521/soco.2012.30.6.689

McGregor, I., Prentice, M., & Nash, K. (2013). Anxious uncertainty and reactive approach motivation (RAM) for religious, idealistic, and lifestyle extremes. *Journal of Social Issues*, *69*(3), 537–563. https://doi.org/10.1111/josi.12028

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology*, *92*(4), 698–716. https://doi.org/10.1037/0022-3514.92.4.698

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. https://doi.org/10.1002/hbm.22466

Musch, J. (2003). Personality differences in hindsight bias. *Memory*, *11*(4–5), 473–489. https://doi.org/10.1080/09658210244000540

Noordewier, M. K., & Breugelmans, S. M. (2013). On the valence of surprise. *Cognition & Emotion*, *27*(7), 1234–1326. https://doi.org/10.1080/02699931.2013.777660

Noordewier, M. K., Topolinski, S., & van Dijk, E. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass*, *10*(3), 136–149. https://doi.org/10.1111/spc3.12242

Park, C. L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin*, *136*(2), 257–301. https://doi.org/10.1037/a0018301

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, *59*(1–2), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X

Piaget, J. (2000). Piaget's theory. In K. Lee & K. Lee (Eds.), *Childhood cognitive development: The essential readings* (pp. 33–47). Blackwell.

Plaks, J. E., Grant, H., & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*(2), 245–262. https://doi.org/10.1037/0022-3514.88.2.245

Pohl, R. F. (2007). Ways to assess hindsight bias. *Social Cognition*, *25*(1), 14–31. https://doi.org/10.1521/soco.2007.25.1.14

Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*(September), 115. https://doi.org/10.3389/fnins.2011.00115

Proulx, T., & Heine, S. J. (2008). The case of the transmogrifying experimenter: Affirmation of a moral schema following implicit change detection. *Psychological Science*, *19*(12), 1294–1300. https://doi.org/10.1111/j.1467-9280.2008.02238.x

Proulx, T., & Inzlicht, M. (2012). The five "A"s of meaning maintenance: Finding meaning in the theories of sense-making. *Psychological Inquiry*, *23*(4), 317–335. https://doi.org/10.1080/1047840X.2012.702372

Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences*, *16*(5), 285–291. https://doi.org/10.1016/j.tics.2012.04.002

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raisig, S., Hagendorf, H., & van der Meer, E. (2012). The role of temporal properties on the detection of temporal violations: Insights from pupillometry. *Cognitive Processing*, *13*(1), 83–91. https://doi.org/10.1007/s10339-011-0413-0

Raisig, S., Welke, T., Hagendorf, H., & van der Meer, E. (2010). I spy with my little eye: Detection of temporal violations in event sequences and the pupillary response. *International Journal of Psychophysiology*, *76*(1), 1–8. https://doi.org/10.1016/j.ijpsycho.2010.01.006

Randles, D., Proulx, T., & Heine, S. J. (2011). Turn-frogs and careful-sweaters: Non-conscious perception of incongruous word pairings provokes fluid compensation. *Journal of Experimental Social Psychology*, *47*(1), 246–249.

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, *7*(5), 411–426. https://doi.org/10.1177/1745691612454303

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

Sleegers, W. W. A., Proulx, T., & van Beest, I. (2015). Extremism reduces conflict arousal and increases values affirmation in response to meaning violations. *Biological Psychology*, *108*, 126–131. https://doi.org/10.1016/j.biopsycho.2015.03.012

Sneed, A. (2014). Fact or fiction? People swallow 8 spiders a year while they sleep. *Scientific American*. https://www.scientificamerican.com/article/fact-or-fiction-people-swallow-8-spiders-a-year-while-they-sleep1/

Stahlberg, D., & Maass, A. (1997). Hindsight bias: Impaired memory or biased reconstruction? *European Review of Social Psychology*, *8*(1), 105–132. https://doi.org/10.1080/14792779643000092

Townsend, S. S. M., Eliezer, D., & Major, B. (2013). The embodiment of meaning violations. In K. D. Markman, T. Proulx, & M. J. Lindberg (Eds.), *The psychology of meaning* (pp. 381–400). American Psychological Association. https://doi.org/10.1037/14040-019

Townsend, S. S. M., Major, B. N., Sawyer, P. J., & Mendes, W. B. (2010). Can the absence of prejudice be more threatening than its presence? It depends on one's worldview. *Journal of Personality and Social Psychology*, *99*(6), 933–947. https://doi.org/10.1037/a0020434

van Maanen, H. (1994). *Encyclopedia of misconceptions*. Boom Uitgevers.

Wyatt, H. J. (1995). The form of the human pupil. *Vision Research*, *35*(14), 2021–2036. https://doi.org/10.1016/0042-6989(94)00268-Q

Zanna, M. P., & Cooper, J. (1974). Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology*, *29*(5), 703–709. https://doi.org/10.1037/h0036651

Author Biographies

Dr. Willem W.A. Sleegers is an Assistant Professor at the Social Psychology department of Tilburg University. His research topics center around false beliefs and how people respond to realizations that their beliefs are false.

Dr. Travis Proulx has worked in collaboration on the Meaning Maintenance Model - a discipline spanning framework that offers an integrated account of inconsistency compensation phenomena. His research focuses on the common ways that people respond to a wide array of meaning violations, ranging from absurdist humor to the absurdity of human mortality.

Prof. Dr. Ilja van Beest is currently a Full Professor of Social Psychology at Tilburg University. His research focusses on social decision making. He seeks to understand the behavioral and physiological consequences of motivated cognition in situations of interdependence.

Handling Editor: Margo Monteith