UC Irvine
UC Irvine Previously Published Works

Title

Misremembering pain: A memory blindness approach to adding a better end.

Permalink

https://escholarship.org/uc/item/96v3v4cn

Journal

Memory & cognition, 47(5)

ISSN

0090-502X

Authors

Urban, Emily J
Cochran, Kevin J
Acevedo, Amanda M
et al.

Publication Date

2019-07-01

DOI

10.3758/s13421-019-00913-9

Peer reviewed

Check for updates

Misremembering pain: A memory blindness approach to adding a better end

Emily J. Urban[1] · Kevin J. Cochran[2] · Amanda M. Acevedo[2] · Marie P. Cross[2] · Sarah D. Pressman[2] ·
Elizabeth F. Loftus[2]

Abstract

How people remember feeling in the past informs future decisions; however, memory for past emotion is subject to a number of biases. Previous research on choice blindness has shown that people often fail to notice when they are exposed to misinformation about their own decisions, preferences, and memories. This type of misinformation can influence how they later remember past events. In the present study, we examined the memory blindness effect in a new domain: memory for pain. Participants ($N = 269$) underwent a cold-pressor task and rated how much pain, distress, and positive and negative affect they had experienced. Later, participants were shown their pain ratings and asked to explain them. Some of the participants were shown lower pain ratings than they had actually made. In a second session, participants recalled how painful the task had been and how much distress and positive and negative affect they remembered experiencing. The results indicated that the majority of participants who were exposed to misinformation failed to detect the manipulation, and subsequently remembered the task as being less painful. The participants in the misinformation condition were not overall more willing to repeat the study tasks, but the participants who recalled less distress, less negative affect, and more positive affect were more willing to repeat the study tasks again in a future experiment. These results demonstrate the malleability of memory for painful experiences and that willingness to repeat aversive experiences may depend more on memory for affective reactions to the original experience than on memory for the physical pain itself.

Keywords Memory · Memory blindness · Pain · Decision making · Emotion

Introduction

When people are asked how willing they would be to do something in the future, like have dinner with their in-laws or schedule a colonoscopy, how do they determine their willingness? One possibility is that they draw on information from their memories of similar past events (Kahneman & Riis, 2005). If previous experiences with one's in-laws have all been positive, one might be excited to have dinner with them. If a prior colonoscopy was painful, one might dread scheduling another and put it off for as long as possible. Future

decision making is guided by memory for emotional responses to relevant past events (Levine, Lench, & Safer, 2009).

If our memories for our past emotional responses guide our future choices, then the accuracy of those memories is highly important. However, both memory and emotion are subject to numerous types of error and bias (Levine et al., 2009; Loftus, 2005; Redelmeier & Kahneman, 1996). For instance, memories for how we felt in the past can be reconstructed and biased by one's present goals and appraisals of past events (Levine et al., 2009), as well as by beliefs about how one typically *should* feel in a specific situation (e.g., McFarland, Ross, & DeCourville, 1989; Robinson & Clore, 2002). Furthermore, memories for emotional events are often biased toward the peak emotional intensity experienced and the emotional intensity experienced at the end of the event (Redelmeier & Kahneman, 1996), and often fail to account for how long the event lasted (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). Thus, our memories for emotionally evocative events can be biased due to natural memory processes.

Emily J. Urban and Kevin J. Cochran contributed equally to this work.

✉ Emily J. Urban
  ejurban@wisc.edu

[1] Center for Healthy Minds, University of Wisconsin, Madison, WI, USA

[2] University of California, Irvine, CA, USA

Despite the fact that memory is subject to bias, people still rely on memories of past emotional experiences to inform their future decisions. One study demonstrated this by exposing participants to each of two different cold-pressor tasks (Kahneman et al., 1993). One of the tasks involved participants holding one hand under circulating water at 14 °C (~ 57 °F) for 60 s; the other task began the same way as the first, and then for 30 s more the temperature of the water increased to 15 °C (59 °F). Participants were then asked which of the two tasks they would rather repeat. One might expect that participants would prefer the shorter version of the task, since it omitted 30 additional seconds of an aversive experience. But the researchers found that more than two thirds of participants thought the longer task caused less discomfort and preferred to repeat the longer task, which included more total pain but a "better end." This study demonstrated that the natural biases existing in memory, particularly for emotional experiences, can lead people to make seemingly illogical decisions.

Misinformation and memory bias

Memories can also be biased due to external influences. For instance, if people are shown misinformation about an event they previously witnessed, their memories about that event are often biased by the misinformation (see Loftus, 2005, for a review). Research using a "false feedback" paradigm has indicated that faulty memories can affect future decision making and behaviors (see Bernstein & Loftus, 2009, for a review). In these studies, participants are typically given a questionnaire assessing their histories eating different foods. Later, they are told their questionnaire was fed into a computer program that suggested they had had certain childhood experiences with particular foods, such as getting sick after eating strawberry ice cream. These studies have demonstrated that many participants come to remember this false event, report decreased preference for the foods implicated, and even eat less of that food when given the opportunity. Conversely, when false feedback is given suggesting something *positive* about a particular food, participants' preference for that food increases. Taken together, these studies demonstrate that *implanted* memories, not just naturally biased memories, can have important consequences for future decision making (Bernstein & Loftus, 2009).

Choice blindness

One method of implanting misinformation is by utilizing the choice blindness procedure. *Choice blindness* occurs when people are misled about their own decisions and evaluations. In a seminal study, participants were shown two pictures depicting different female faces and asked to select the one they found more attractive (Johansson, Hall, Sikström, & Olsson, 2005). After making their decision, participants were given the photograph they selected and asked to justify why

they made the choice they did. However, on certain trials, a sleight-of-hand manipulation took place whereby the photograph participants were given was the nonselected photograph instead of the one they had selected. Not only did a majority of people fail to detect when their decisions had been manipulated, but they generated reasons why they made decisions they had never truly made. Choice blindness has been examined in a variety of contexts, including financial decision making (McLaughlin & Somerville, 2013), taste-testing of grocery products (Hall, Johansson, Tärning, Sikström, & Deutgen, 2010), and even one's own reported history of criminal and norm-violating behavior (Sauerland et al., 2013).

Memory blindness

The choice blindness paradigm, which is typically tested in the context of short-term attitude change, has recently been extended into studies on memory, in which researchers investigate whether choice blindness can lead to memory change (Cochran, Greenspan, Bogart, & Loftus, 2016; Stille, Norin, & Sikström, 2017). In one such experiment, participants watched a simulated crime and were then asked to answer questions about the event, such as "how tall was the thief?" (Cochran et al., 2016). Later, they were shown their responses to these questions and asked additional follow-up questions. However, for misinformation items, participants' responses had been manipulated (e.g., the reported height of the thief was increased or decreased). Participants were asked to report their memories for the initial event again later in the session. The researchers found that for misinformation items, participants' memories shifted in the direction of the misinformation they received, whereas participants' memories did not shift for control items. In a second study, participants watched a simulated crime and were asked to identify the perpetrator from a lineup. After receiving misinformation suggesting that they had identified a different person than they actually had, participants were more likely to switch their identification in a subsequent lineup. These studies demonstrate the downstream consequences of choice blindness: not only do people often fail to detect misinformation about their own memory reports, their subsequent memories can become biased by the misinformation they receive.

Misinformation and healthcare

Healthcare settings might be one context in which misinformation could be especially consequential. Patients are often asked by medical professionals to describe their physical and psychological symptoms as well as their levels of pain and discomfort. People may be susceptible to remembering their symptoms or pain differently as a result of misinformation, which could then influence the healthcare decisions they make in the future. On the other hand, pain might be less amenable to misinformation than are other affective experiences, given

the salience of pain in the moment and thus greater attention to the details of the experience (Eccleston & Crombez, 1999). Therefore, it is unclear whether memory for pain could be altered by misinformation in the same way it is susceptible to natural memory biases (e.g., peak and end bias; Redelmeier & Kahneman, 1996).

A handful of studies have utilized misinformation in the context of psychological and physical healthcare. One study employed false feedback to influence peoples' overall memories for painful, stressful, and uncomfortable procedures. In this study, the researchers examined children who had received their diphtheria pertussis tetanus shots (Bruck, Ceci, Francoeur, & Barr, 1995). Approximately 11 months after the inoculations, the children participated in three interviews in which they received either neutral or pain-denying feedback (i.e., feedback that the shot did not hurt). The participants who received the pain-denying feedback remembered less pain and also that they had cried less than those who received neutral feedback.

Another study misled participants about the frequency with which they reported experiencing psychological symptoms, such as repeated unpleasant thoughts (Merckelbach, Jelicic, & Pieters, 2011). Participants reported their symptoms using a 0–4 scale, where 0 indicated *not at all* and 4 indicated *all the time*. Later, participants were shown their responses to some of the items and were asked to recall why they gave those ratings. However, the researchers surreptitiously increased participants' ratings on two items by two scale points. Participants were then given the questionnaire a second time for an immediate retest, and were given the questionnaire a third time one week later.

The researchers found that 63% of participants were unaware of the manipulation. Furthermore, whereas these "blind" participants did not differ in their ratings of manipulated and control symptoms at baseline, they rated the manipulated symptoms significantly higher at both immediate and one-week follow-ups. Nonblind participants showed no difference between manipulated and control symptoms at any time. A more recent article replicated these findings using a symptom checklist that included both psychological and somatic symptoms and demonstrated that participants could also be led to underestimate their symptom ratings as a result of misinformation (Merckelbach, Dalsklev, Van Helvoort, Boskovic, & Otgaar, 2018). These studies illustrate that people can be misinformed about their own internal states. Moreover, this misinformation causes people to report feeling differently; if they are told they reported having more unpleasant thoughts, they actually report experiencing more unpleasant thoughts.

The aforementioned studies examined whether misinformation, and more specifically memory blindness, could be used to change memory for physical and psychological symptoms. To our knowledge, no study has examined memory blindness for physical pain ratings among adults, nor how memory blindness in a health relevant setting might be used to make health-related decisions in the future. One potential

application of using memory blindness in a medical setting is to increase compliance for routine, yet mildly painful, medical procedures. If patients recall pain experienced in the medical setting as less painful than they originally reported, they may be more willing to seek out medical care in the future. Leveraging memory bias to increase compliance for routine medical procedures is not necessarily novel. One study used the principle of duration neglect to increase the odds that patients would return for a repeat colonoscopy by subjecting them to a longer initial colonoscopy (but ended with a period of less intense pain; Redelmeier, Katz, & Kahneman, 2003). Although this study was successful at increasing medical compliance, memory blindness provides a potential avenue to alter memory for painful experiences without extending the duration of the pain.

The present study

The goal of the present study was twofold: (1) to extend the memory blindness literature to the domain of physical pain and examine whether memory blindness can be used to reduce remembered pain experienced during a painful laboratory task, and (2) to examine how memory for pain (whether accurate or retrospectively biased) might influence intentions for future behavior as measured through willingness to repeat the painful task.

Participants were recruited for a two-session study. In the first session, participants underwent a cold pressor task (Mitchell, MacDonald, & Brodie, 2004). The cold pressor is a well-established pain induction technique, because the discomforts are relatively short lived (the task only lasts a few minutes) and normal sensation is recovered rapidly. In the stress reactivity literature, it is considered a noninvasive method of inducing a cardiovascular response (Goyal, Shimbo, Mostofsky, & Gerin, 2008). Immediately following the task, participants rated how painful the experience was on a 100-point scale. They also rated how much distress and positive and negative affect they felt during the task. Later in the first session, participants were reminded of the pain rating they had produced and asked to elaborate on why they had rated their pain the way they did. However, unbeknownst to the participants, those in the misinformation group were told they had rated their pain 20 points lower than they truly had. Participants returned to the lab one to two days later and were asked to recall how painful the task was and how much distress and positive and negative affect they had experienced, to rate how willing they would be to participate in a similar experiment again in the future, and to recommend how much money participants should be compensated in a similar experiment in the future.

We expected that the majority of participants in the misinformation condition would fail to detect the misinformation about their pain ratings. We hypothesized that the difference

between recalled and experienced pain would be larger for the participants exposed to the misinformation than for controls, particularly for those who failed to detect the misinformation. We also hypothesized that the more pain that the participants in both conditions remembered from the task, the less willing they would be to repeat the experience and the more compensation they would recommend paying participants for a future study. Since the participants in the misinformation condition were given misinformation indicating that they had reported experiencing *less* pain, we expected them to recall less pain and to report being more willing to undergo a similar study in the future than controls would be. However, we predicted that these effects would be attenuated for those participants who detected the misinformation.

Supplementary analyses were also performed to test whether misinformation regarding pain ratings influenced recollections of distress, positive affect, and negative affect. Although these relationships were not of primary interest, we wanted to test whether the misinformation manipulation had more localized (i.e., constrained only to pain) or more global (i.e., transferring to affective responses related to pain) effects. Because experiences of pain often produce affective responses, we predicted that underestimating pain due to the misinformation would be related to also underestimating distress and negative affect and overestimating positive affect. Regardless of the effect of the pain misinformation on recalled distress, positive affect, and negative affect, we predicted that participants would be more willing to repeat the study tasks, the less distress, the less negative affect, and the more positive affect they had experienced during the task and recalled during the follow-up.

Method

Participants

The overall sample size was determined on the basis of experience with previous choice blindness research (Cochran et al., 2016). We wanted to recruit a sufficient sample to observe many detectors and nondetectors, in the case that our manipulation had either a high or low rate of detection. A total of 303 eligible participants were recruited from the university's human participant research pool in exchange for course credit. Participants between the ages of 18 and 35 were eligible for the study but were screened out if they had an injury to their nondominant hand, had Reynaud's disease or a similar circulatory problem, had vasovagal syncope, had ever fainted or had a seizure, were pregnant, or regularly took mood-altering, pain-altering, or cardiovascular-functioning medication. These exclusions were used for safety reasons and for their potential influence on emotional responses to the cold pressor. A subset of participants were excluded from the analyses

because of incomplete data ($n = 10$) or equipment malfunction ($n = 24$).

The final sample of 269 participants (control group $n = 134$, misinformation group $n = 135$) was 82.5% female, 17.5% male, and 50.2% Asian/Pacific Islander, 31.6% Hispanic/Latino, 8.9% White, 1.5% African American, 4.8% biracial, and 3% other race/ethnicity. The mean age of the participants was 21 years old ($SD = 2.32$, range = 18–33). Of these participants, 244 completed the second study session; 233 (95.5%) of these did so one or two days following their initial session, and these participants composed the final sample (control group $n = 111$, misinformation group $n = 122$). The other 11 participants completed the second session either on the same day as the first session or three or more days following the first session, and are not included in analyses on the data collected during Session 2.

Design and procedure

Session 1

All study procedures were approved by the university's Institutional Review Board. Participants were first given details about the study, including the cold pressor. After agreeing to participate, they completed an eligibility questionnaire. Those who did not meet the eligibility criteria were dismissed and received partial course credit. If participants met the eligibility criteria, they completed a series of baseline questionnaires, including questions on emotional states, while the cold pressor was being prepared.

Cold-pressor task After the baseline period was completed, the research assistant brought in a bucket filled with water measuring 4 °C (± 0.4 °C) with a thermometer attached to the bucket out of the participant's sight. The bucket was also affixed to a motorized water-circulating pump to ensure constant flow of water around the participant's hand. The research assistant then instructed participants to place their nondominant hand in the bucket of cold water for 90 s. Participants were told they could remove their hand if the task became too uncomfortable and they felt they were unable to finish.

Recovery period Immediately after the participant removed his or her hand from the water bucket, a 5-min recovery period began. Participants rated on 100-point sliding scales how painful the task had been and how much distress, positive affect, and negative affect they had experienced during the task. After they finished these questionnaires, they were asked to sit quietly until the resting period was over.

Misinformation manipulation Once the resting period was completed, participants were instructed to complete a final set of questionnaires that included a prompt reminding them of the

pain score they had made immediately following the cold pressor. Participants were asked, "We are interested in understanding what the experience of putting your hand in cold water was like for you. Earlier in the study, on a scale from 1 to 100,[1] you rated your pain as a ____. What made you rate the task in that way? Please be detailed in your response." The computer program randomly assigned participants to either the control condition or the misinformation condition. In the control condition, participants were reminded of the actual pain score they had reported earlier. In the misinformation condition, participants were told they had given a pain score that was 20 points lower than the one they had reported earlier. Once all questionnaires were completed, participants were partially debriefed (the fact that pain ratings had been altered for half the participants was withheld) and dismissed from the study.

Session 2

One to two days following the first study session ($M = 1.28$ days, $SD = 0.49$), participants returned and completed a series of questionnaires. Participants were asked to report again how much pain, distress, positive affect, and negative affect they had experienced during the cold pressor, using the same wording and 100-point sliding scale from the first session. At a predetermined time between questionnaires, after all questions regarding memory for the cold pressor, the research assistant interrupted participants to ask them to fill out a paper-based survey that was ostensibly unrelated to the present study (see the Measures section). After the survey was completed, participants finished the set of questionnaires on the computer and were debriefed. Part of the debriefing portion included questions assessing whether participants had detected the misinformation since the first session.

Measures

Pain, distress, and overall positive and negative affect

Immediately following the cold pressor, participants were asked to rate how much pain, distress, and positive and negative affect[2] they had felt during the cold pressor, on sliding scales of 0 *No pain/distress/positive emotion/negative emotion* to 100 *Most pain/distress/positive emotion/negative emotion imaginable*. At follow-up, participants were asked these same questions regarding pain, distress, and affect again, on the same scale. The items regarding distress and affect were

included to test whether misinformation regarding the pain reports might influence the recalled distress and affect, or whether the manipulation would exclusively lower recalled pain in memory.

Social desirability

The Marlowe–Crowne Social Desirability Scale (Crowne & Marlow, 1960), which includes 33 questions that are answered either "true" or "false," was used to measure the extent to which participants tended to present themselves in socially desirable ways. An example item is "My table manners at home are as good as when I eat out in a restaurant." The socially desirable responses endorsed (15 of which are reverse-coded) are then added up to create a total social desirability score, which ranges from 0 to 33. We used the scores from this scale to measure the potential influence of demand effects on pain ratings and the tendency to report detection of the misinformation. People who have high social desirability may be more likely *not* to report that they detected misinformation (even if they did detect it) and may also report lower pain, distress, and negative affect ratings and higher positive affect and willingness ratings.

Concurrent and retrospective detection of misinformation

On the basis of prior choice and memory blindness research, two measures were created to record whether participants detected that their pain ratings had been manipulated. To assess whether participants detected the misinformation concurrently (i.e., in response to the presentation of the misinformation), three trained coders evaluated the participants' typed responses to the open-ended question asking them to elaborate on why they had given that specific pain rating after being reminded of the pain rating. If two of the three trained coders flagged a response as revealing suspicion that the number they were being shown was not actually their true rating, the participant was coded as a concurrent detector. For example, if a participant typed, "that rating is lower than the one I gave, the water was actually pretty painful," that participant would be labeled as a concurrent detector.

A second measure of detection was created based on the responses provided in the second session (retrospective detection of misinformation). The labeling of this variable as "retrospective" is intended to reflect when participants *revealed* that they had noticed the misinformation, rather than when they had actually noticed it. That is, it is possible that participants could have noticed the misinformation during the first session (concurrently) but not revealed their suspicions until the second session (retrospectively). Participants were asked during the debriefing whether they remembered being reminded of their pain rating score during the first session and being asked to write about their pain experience. The research assistant, who was unaware of the participant's condition, then asked, "Did you notice anything strange about this process?"

---

[1] The scale on which participants reported their pain ranged from 0 to 100, but due to a clerical error, the critical manipulation read "1 to 100." It was not expected that this would have any influence on the results of the present study.
[2] Questions regarding positive and negative affect consisted of one item each. The participant saw the word "emotion" in place of the word "affect," but since these questions were regarding general positive or negative states, we refer to these measures as *positive* and *negative affect* throughout the article.

The research assistant recorded the participant's verbal response in the study notes, which was later coded by trained coders. Participants were coded as retrospective detectors if they mentioned finding anything strange about the process (e.g., "I noticed the computer gave me a lower pain rating"), unless what they found strange was irrelevant to the study's hypotheses (e.g., "I was surprised by how many questionnaires there were"). Three research assistants independently coded each participant's response to these questions, and participants were labeled as detectors if at least two of the research assistants coded them as such.

Willingness to participate in a similar future task

During the second session, participants were given a paper-based survey that was ostensibly designed to measure their opinions regarding the study, so that researchers would know how much to pay future participants in a similar study. The questions that were asked were instead intended to measure how willing the participant was to repeat the cold-pressor experience. The survey included the questions "How willing would you be to participate in an experiment similar to this one in the future?" on a scale of 1 *Not at all willing* to 5 *Extremely willing*, and, "How much do you think we should pay our participants?," which was open-ended. Because these items did not distinguish between the whole study and just the cold-pressor portion (which was the portion that the main manipulation targeted), we altered these two questions for the final 49 participants[3] to specify their relevance to completing the entire study, including the cold pressor. The same subset of participants also saw two additional questions asking how willing they would be to do just the cold-water task, on a scale of 1 *Not at all willing* to 5 *Extremely willing* (asked before the willingness question regarding the entire study) and how much future participants should be paid to do just the cold-water portion of the study, on a scale from $0 to $60 (in $5 increments, asked before the compensation question regarding the entire study).

Results

In cases in which the assumption of homogeneity of variances was violated, corrected statistics are presented. This is particularly relevant for analyses concerning detection status, where unequal cell sizes make the violation of this assumption more likely. For the analyses relating to our main hypotheses, we report the Bayes factor in addition to listing the effect size and level of significance. Bayes factors (BFs) allow for a comparison of the likelihood that the observed data fit with the null

hypothesis (H_0), as compared to the likelihood that the data fit with the alternative hypothesis (H_1; see Wagenmakers, Marsman, et al., 2018b). The statistical software JASP, version 0.9, was used to calculate BFs (JASP Team, 2018). A BF_{10} refers to the relative likelihood that the observed data fit the alternative hypothesis (represented by the subscript 1) as compared to the null hypothesis (represented by the subscript 0). Whereas a BF_{10} of less than 1 indicates more evidence in support of H_0, a BF_{10} of more than 1 indicates some degree of evidence in support of H_1 (Wagenmakers, Love, et al., 2018a). BF_{10}s greater than 3 are considered at least moderate evidence for H_1 (1–3 is considered anecdotal evidence; 10–30 is considered strong, 30–100 is considered very strong, and greater than 100 is considered extreme evidence in support of H_1).

Descriptive statistics

A total of 179 participants[4] (66.5%) completed the entire 90 s cold-pressor period. The average length of time for those who removed their hand early was 36.19 s (SD = 19.11 s). A significant Shapiro–Wilk test ($p < .001$) demonstrated that the number of seconds that participants kept their hand in the water was significantly skewed (– 1.09, SE = 0.15); thus, nonparametric tests were conducted to examine any relationships with this variable. Mann–Whitney U tests revealed that men and women did not differ significantly in how long they kept their hand in the water (p = .113), and that cold-pressor duration did not vary as a function of condition (p = .513). Using Spearman's correlations, we found that the longer participants that kept their hand in the water, the less pain they reported experiencing during the cold pressor, r_s (268) = – .13, p = .031. The amount of time they kept their hand submerged in the water was not related to distress, positive affect, or negative affect during the cold pressor, ps > .08. Participants in misinformation conditions did not differ significantly in gender composition, $\chi^2(1)$ = 0.04, p = .850, or reported pain, distress, positive affect, or negative affect experienced during the cold pressor, ps > .32. See Table 1 for descriptive statistics regarding reported and recalled pain, distress, positive affect, and negative affect as a function of condition.

Social desirability

Average levels of social desirability were similar across conditions, $t(231)$ = – 0.98, p = .327. Social desirability was not related to the amount of time participants kept their hand in the water, their reported pain, or their reported distress during the cold pressor, ps > .08. People who detected the misinformation concurrently scored lower on the measure of social

---

[3] The question regarding payment for the entire study for these 49 participants was on a scale of $0 to $60 (in $5 increments).

[4] Due to experimenter error, a total of 93 s elapsed before one participant was instructed to remove their hand. This participant was not included when using seconds in water as a continuous variable, but was included in analyses using whether or not they removed their hand before 90 s.

Table 1 Experienced and recalled pain, distress, negative affect, and positive affect by condition

	Pain	Distress	Negative Affect	Positive Affect
	Experienced M (SD)			
Control	67.30 (19.48)	62.58 (22.41)	56.19 (24.47)	16.96 (18.31)
Misinformation	68.70 (20.23)	61.13 (23.28)	57.59 (26.67)	16.20 (19.31)
	Recalled M (SD)			
Control	64.38 (19.84)	61.64 (22.82)	56.01 (24.54)	18.75 (19.23)
Misinformation	57.79 (21.89)	56.24 (25.38)	54.29 (25.93)	20.91 (21.85)

Control group: $n = 134$ for experienced ratings, $n = 111$ for recalled ratings; misinformation group: $n = 135$ for experienced ratings, $n = 122$ for recalled ratings.

desirability than did those who did not detect the misinformation, $t(120) = 3.09$, $p = .002$, as did people who detected the misinformation retrospectively, $t(120) = 1.98$, $p = .05$, although this difference for the retrospective detectors did not reach significance.

Higher levels of social desirability were related to higher positive affect and lower negative affect during the cold pressor, rs > |.13|, ps < .033. Furthermore, levels of social desirability were related to recalling less pain, less distress, less negative affect, and more positive affect during the second study session, rs > |.21|, ps < .002. People higher in social desirability were also more willing to participate in a future research study, $r(231) = .14$, $p = .038$, but social desirability scores were not related to suggested payment for future participants, rs < .18, ps > .257.

Because social desirability was related to recalling less pain, we ran a correlation among the participants in the misinformation condition to explore whether people with higher levels of social desirability were more susceptible to the misinformation. Higher levels of social desirability were related to a greater underestimation of pain at Session 2, $r(120) = -.32$, $p < .001$, suggesting that the participants with higher levels of social desirability were more susceptible to the misinformation.

Detection of misinformation

Coder agreement for classifying whether participants were concurrent or retrospective detectors was high (all average pairwise percent agreements ≥ 95%, all Krippendorff's alphas ≥ .75). Of the participants in the misinformation condition, only 14 (11%) detected the misinformation concurrently. No participant in the control condition was judged to be a concurrent detector. For retrospective detection, 42 (34.4%) of the participants in the misinformation condition were coded as detectors, whereas seven control participants (6.3%) were coded as detectors. For the subsequent analyses, only the participants in the misinformation condition who were coded as detectors were considered detectors; the control condition was not partitioned.

It should be emphasized that detection status is a quasi-independent variable that characterizes differences between misinformation participants but cannot be randomly assigned or experimentally manipulated. Therefore, any relationship between detection status and a dependent variable reported here is correlational in nature. Furthermore, any conclusions drawn from analyses with small cell means (as in the case of concurrent detectors) should be interpreted with caution and replicated with a larger sample size.

Participants who detected the misinformation concurrently reported marginally more pain, $t(133) = -1.98$, $p = .05$, and significantly more distress, $t(24.55) = -3.87$, $p = .001$, and negative affect, $t(133) = -3.23$, $p = .002$, during the cold pressor than did those who did not detect the misinformation concurrently. As compared to nondetectors, a greater proportion of the concurrent detectors took their hand out of the water before 90 s had elapsed, $\chi^2(1) = 7.01$, $p = .008$. Detection status did not vary as a function of gender, $\chi^2(1) = 0.16$, $p = .686$.

Participants who detected the misinformation retrospectively did not differ from nondetectors in levels of reported pain, $t(120) = -0.40$, $p = .691$, or distress, $t(120) = -1.66$, $p = .099$, but they did report significantly higher negative affect, $t(120) = -2.78$, $p = .006$, and significantly lower positive affect, $t(118.91) = 2.85$, $p = .005$, during the task. A similar proportion of participants pulled their hand out of the water early, regardless of retrospective detection status, $\chi^2(1) = 0.06$, $p = .80$, and retrospective detection status did not vary as a function of gender, $\chi^2(1) = 1.45$, $p = .229$.

Memory bias for pain

To assess bias in participants' memory for the cold pressor, a mixed analysis of variance (ANOVA) was computed with the pain rating type entered as the within-participants variable (experienced pain vs. recalled pain) and condition entered as the between-participants variable (control vs. misinformation). No main effect of condition was found, $F(1, 231) = 1.00$, $p = .319$, $\eta_p^2 = .004$. A main effect of rating type was found, $F(1, 231) = 128.21$, $p < .001$, $\eta_p^2 = .36$, as was a
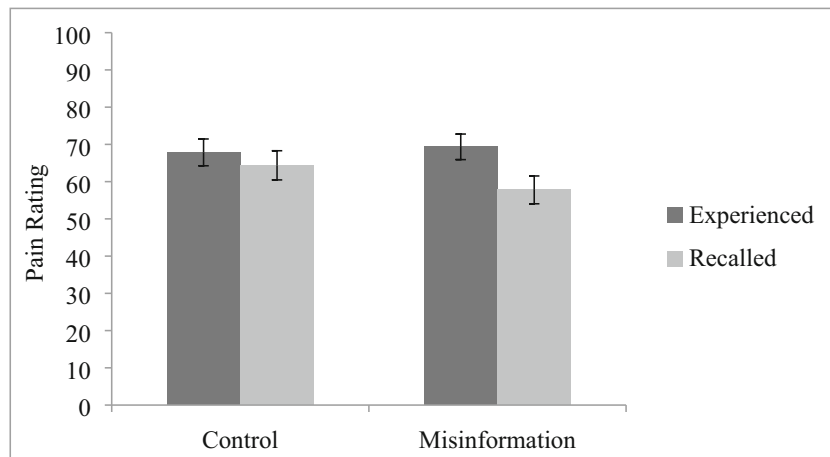
Fig. 1. Experienced and recalled pain by condition. Error bars represent 95% confidence intervals. Control group $n = 111$, misinformation group $n = 122$

significant interaction between condition and pain rating type, $F(1, 231) = 37.03$, $p < .001$, $\eta_p^2 = .14$. The model including the main effects of condition and rating type as well as their interaction had the highest BF_{10}, of 1.18e25, indicating that including the interaction in the model provided the best fit for the data. Overall, participants tended to underestimate how painful the cold pressor was in their recollections, but the participants in the misinformation condition underestimated their pain during recall ($M_{diff} = -11.56$) to a greater degree than did those in the control condition ($M_{diff} = -3.48$). A paired t test revealed that although the participants in the control condition underestimated their pain to a lesser degree, their recalled pain was still significantly lower than their reported pain, $t(110) = 3.85$, $p < .001$, $BF_{10} = 86.51$. See Fig. 1 for a graphical representation of this relationship.

To determine whether concurrent detection of the misinformation was associated with participants' bias in their memories for their pain, a 3 (detection status: control vs. nondetectors vs. detectors) by 2 (report type: experience vs. memory) mixed ANOVA was conducted.

This analysis revealed a main effect of concurrent detector group status, $F(2, 230) = 5.25$, $p = .006$, $\eta_p^2 = .04$; a main effect of report type, $F(1, 230) = 31.29$, $p < .001$, $\eta_p^2 = .12$; and a significant interaction between concurrent detector group status and report type, $F(2, 230) = 32.60$, $p < .001$, $\eta_p^2 = .22$. The BF_{10} for the model that included the main effects of condition and rating type as well as their interaction was the largest, at 1.20e30. A follow-up one-way ANOVA and a Games–Howell post-hoc test for multiple comparisons revealed that the degrees of memory bias were significantly different across all three groups, ps $< .048$. As can be seen in Fig. 2, the participants in the misinformation condition who failed to detect the misinformation concurrently showed the greatest memory bias for pain ($M_{diff} = -13.11$), followed by the participants in the control condition ($M_{diff} = -3.48$). A series of paired-sample t tests showed that the degrees of underestimation were significant for both the control condition and the group that did not detect the
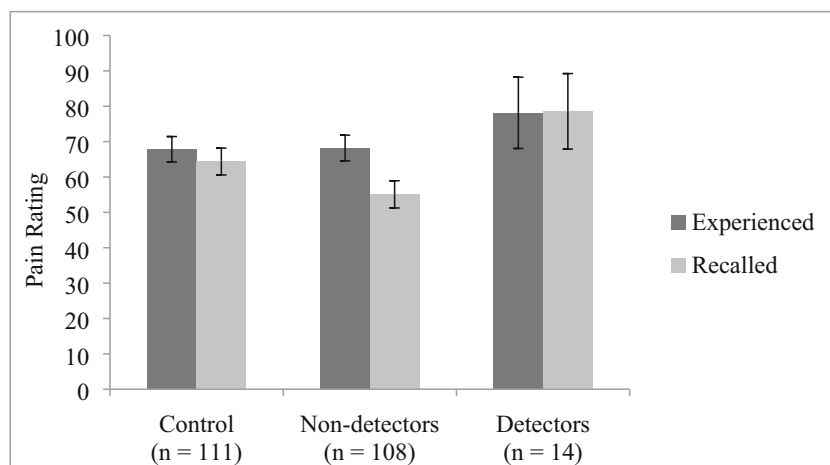

Fig. 2. Experienced and recalled pain by concurrent detection status. Error bars represent 95% confidence intervals
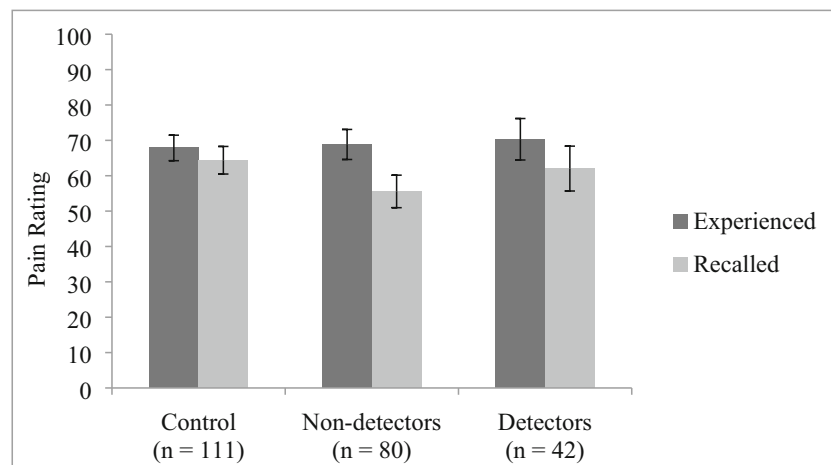
Fig. 3. Experienced and recalled pain by retrospective detection status. Error bars represent 95% confidence intervals

misinformation, $ts > 3.81$, $ps < .001$, $BF_{10} > 86.50$. Participants in the misinformation condition who detected the misinformation concurrently, however, showed no significant memory bias ($M_{diff} = 0.43$), $t(13) = -0.34$, $p = .74$, $BF_{10} = 0.28$.

A similar 3×2 mixed ANOVA was computed to determine whether retrospective detection status influenced the bias participants exhibited in their memories for their pain. The results indicated no main effect of retrospective group detection status, $F(2, 230) = 1.07$, $p = .34$, $\eta_p^2 = .01$. There was a significant main effect of pain rating time point, $F(1, 230) = 138.42$, $p < .001$, $\eta_p^2 = .38$, which was qualified by an interaction with retrospective group detection status, $F(2, 230) = 22.41$, $p < .001$, $\eta_p^2 = .16$. As with the previous two mixed ANOVAs, the BF_{10} was largest for the model incorporating both main effects and the interaction of detection status and rating time point, with a BF_{10} of 4.80e25. Overall, participants tended to underestimate how much pain they had experienced when asked at recall, but a follow-up one-way ANOVA and Bonferroni-corrected post-hoc comparison tests revealed that the degrees of bias were significantly different across all three groups, $ps < .029$. Follow-up paired-sample t tests revealed that all three groups tended to underestimate their pain significantly, $ts > 3.82$, $ps < .001$, $BF_{10} > 86.50$, but memory bias for pain was the greatest for participants who did not detect the misinformation retrospectively ($M_{diff} = -13.28$). Those who detected the misinformation retrospectively still demonstrated a significant memory bias ($M_{diff} = -8.29$), which was greater than the bias shown by the participants in the control condition ($M_{diff} = -3.48$). Together, the results of these analyses suggest that people tended to recall the cold pressor as being less painful than they had initially reported it as being one to two days earlier, but this underestimation was magnified for those in the misinformation condition,

particularly for those who did not detect the misinformation retrospectively (see Fig. 3).

Willingness to participate in the future

We computed correlations between recalled pain and our four measures of willingness (willingness for the entire study, willingness for just the cold pressor, recommended payment for the entire study, and recommended payment for just the cold pressor; see Table 2 for descriptive statistics broken down by condition).[5] As can be seen in Table 3, we found one significant, yet weak, relationship: The more pain that participants recalled, the more compensation they recommended for future participants in the overall study.[6] There was no significant relationship between recalled pain and any of the other three willingness measures. We also found that pain experienced *during* the cold pressor was not related to any of the willingness variables. Given that willingness only varied as a function of recalled pain for recommended overall compensation, we tested for condition differences in willingness using this variable only. We found no differences between the misinformation and control groups in how much compensation they recommended future participants receive, regardless of

---

[5] Independent-samples t tests were conducted to test for differences in responses as a result of changes in the question wording. These tests revealed no significant differences in willingness as a function of the old versus new wording for the whole sample, for the misinformation group, or for controls, $ps < .09$. The amount recommended that future participants be paid did vary as a function of the old versus new wording for the whole sample and for the misinformation group, $ps < .02$, but not for controls, $p = .539$. Given these findings, the first willingness question was combined, regardless of the wording change (for a total $n = 233$), but suggested payment was analyzed separately, depending on whether participants were given the open-ended ($n = 184$) or scale ($n = 49$) question format. Analyses regarding willingness and payment specifically for the cold pressor included $n = 49$.

[6] This analysis was performed with the participants who recommended compensation in the open-ended question ($n = 184$). The same analysis for participants who recommended compensation in the interval-scale question ($n = 49$) was nonsignificant, $r(47) = -.14$, $p = .339$.

Table 2 Descriptive statistics of the willingness and compensation variables by condition

	Whole-Study Willingness M (SD)			Cold-Pressor Willingness M (SD)	
	Willingness[a]	Compensation (Open-ended)[b]	Compensation (13-point scale)[c]	Willingness[c]	Compensation[c]
Control group	3.95 (0.89)	15.27 (14.09)	17.20 (12.51)	3.56 (1.00)	13.20 (9.45)
Misinformation group	4.02 (0.93)	12.44 (8.36)	21.67 (17.30)	3.58 (1.21)	15.73 (12.54)

Due to changes in the willingness questionnaire mid-study, sample size for each measure differ. For control group: [a] $n = 111$. [b] $n = 86$. [c] $n = 25$. For misinformation group: [a] $n = 122$. [b] $n = 98$. [c] $n = 24$.

whether the question was asked in open-ended or scale format, ts < 1.64, ps > .10, BF_{10} < 0.61.

Predicting willingness for future behavior

Since pain was not related to willingness to participate in a similar future study, we conducted a series of exploratory correlational analyses to assess whether distress, positive affect, or negative affect might be related to willingness (see Table 3). We found that participants were more willing to complete a similar future study if they recalled less distress, less negative emotion, and more positive emotion.[7] Experienced distress and experienced negative affect reported during the cold pressor, however, were not significantly related to willingness. Positive affect experienced during the cold pressor was related to greater willingness to participate in a similar future study.[8] These analyses should be interpreted with caution, however, due to their exploratory nature and the inflated risk for Type I errors, given the presence of multiple comparisons.

Memory bias for distress, positive affect, and negative affect

We explored whether misinforming participants of their pain rating carried over to their memory for the amounts of distress, positive affect, and negative affect they experienced during the cold pressor. To test this, three mixed ANOVAs were computed with condition as the between-participants variable (control vs. misinformation) and report type as the within-participants variable (experienced vs. recalled) for distress, positive affect, and negative affect separately. In all three ANOVAs, no main effect was found for condition, Fs < 1.54, ps > .21, but the main effect for report type was significant, Fs > 4.82, ps < .03. Thus, participants tended to underestimate the distress and negative affect they had experienced when later recalling how they had felt during the cold pressor, and they tended to overestimate how much positive affect they

had experienced. Contrary to our hypothesis, the interaction between condition and reporting session was not significant for any of the three models, ps > .09, demonstrating that misinformation regarding their pain rating did not influence participants' memory for their emotional reactions to the cold pressor.

Discussion

This study demonstrated that people can be misled about their own reports of the pain they experienced from a cold pressor. Participants who received misinformation regarding their reported pain later exhibited a greater memory bias (i.e., underestimated their pain rating to a greater extent) than did control participants who did not receive misinformation. This effect was amplified for participants who failed to detect that they had been given misinformation about their pain ratings. Participants who retrospectively detected the misinformation exhibited a greater reduction in their pain ratings than did control participants, but a lesser reduction than participants who failed to detect the misinformation retrospectively. However, participants who concurrently detected the misinformation did not exhibit a reduction in their pain ratings. These findings are consistent with past research demonstrating that people can be led to misremember their own reports on their internal states (Merckelbach et al., 2018), that choice blindness can have lasting effects for memory (i.e., memory blindness; Cochran et al., 2016; Stille et al., 2017), and that when people detect the discrepancy between misinformation and facts, they are less likely to be swayed by the misinformation (Tousignant, Hall, & Loftus, 1986). These findings add to the literature by demonstrating that memory blindness can be found in memory for a painful, lived experience, not just in symptoms on a checklist.

This study also examined the influence of biased memory for pain on intentions for future behavior. Memories for past experiences are used to inform decisions made in similar situations in the future (Levine et al., 2009). Despite this, in the present study we found only weak evidence that remembered pain was used to inform willingness to repeat the painful experience in the future (recalled pain was weakly related to

---

[7] When running these correlations separately by condition, the results were similar, except that the relationship between willingness and recalled negative emotion among control participants was no longer significant, $r(111) = -.16$, $p = .094$, $BF_{10} = 0.47$.

[8] But this was not the case when running the correlation among controls only, $r(111) = .14$, $p = .143$, $BF_{10} = 0.34$.

Table 3 Correlations of experienced and recalled pain with the willingness variables

| | | Whole-Study Willingness | | | | | | Cold-Pressor Willingness | | | |
| | | Willingness[a] | | Compensation (Open-ended)[b] | | Compensation (13-point scale)[c] | | Willingness[c] | | Compensation[c] | |
		r	BF_{10}	r	BF_{10}	r	BF_{10}	r	BF_{10}	r	BF_{10}
Experienced	Pain	− .07	0.14	.11	0.30	− .02	0.18	− .05	0.19	.03	0.18
	Distress	− .10	0.25	.13	0.42	− .06	0.19	.01	0.18	− .04	0.19
	Positive affect	.18[**]	3.27	− .02	0.09	− .22	0.52	.23	0.64	.03	0.19
	Negative affect	− .07	0.13	.10	0.24	− .06	0.19	.10	0.23	− .10	0.23
Recalled	Pain	− .11	0.30	.15[*]	0.78	− .14	0.28	− .08	0.20	− .02	0.18
	Distress	− .21[**]	14.58	.11	0.26	− .09	0.21	− .07	0.20	− .08	0.21
	Positive affect	.27[**]	399.57	− .06	0.12	− .26	0.85	.26	0.91	− .08	0.21
	Negative affect	− .17[**]	2.51	.09	0.19	.03	0.18	− .02	0.18	.04	0.19

Due to changes in the willingness questionnaire mid-study, sample size for each measure differ. [a] $n = 233$. [b] $n = 184$. [c] $n = 49$. [**] $p < .01$; [*] $p < .05$.

suggesting less compensation for future participants when the question was asked in an open-ended format). Instead, exploratory analyses revealed that memory for affective experiences related to the pain, such as distress, negative affect, and positive affect, might instead be more influential on behavioral intentions to repeat painful tasks. Replication of these findings is warranted, as is further research to determine the role played by affective memory biases in the willingness to repeat painful experiences.

Memory blindness for pain

Past research has shown that pain is susceptible to naturally occurring memory biases (Kahneman et al., 1993; Redelmeier & Kahneman, 1996). Because of the attention-grabbing nature of pain (Eccleston & Crombez, 1999), it is reasonable to believe that memories of pain might be less amenable to the influence of misinformation. Contrary to this intuition, the present study demonstrated that participants in the misinformation condition exhibited a greater decrease in their memory for pain than did those in the control condition, particularly when they did not detect the misinformation. It seems, then, that pain is not different from the typical targets of memory blindness studies, in that memory of pain is indeed susceptible to external influences. There may be a limit on the extent of this susceptibility, however, since the participants in the misinformation condition were less susceptible to underestimating their pain levels the more pain they had initially reported during the task.

Predicting willingness for similar future experiences

Contrary to our hypothesis, recalled pain was related to only one of four indices intended to measure willingness for undergoing a similar lab experience in the future. The single significant, yet weak, correlation revealed that the more pain participants recalled, the more money they recommended future participants be compensated for the entire experiment. Recalled pain was not related to any other index of willingness, and experienced pain was not related to any of the willingness indices. Furthermore, we found no differences in willingness between the misinformation and control condition, which were designed to vary solely as a function of recalled pain. These findings together suggest that participants did not rely on remembered pain when making their decisions about willingness to undergo a similar experience in the future. Although initially counterintuitive, this may make sense in the context of past research in which participants preferred the study task that involved more overall pain (by choosing the longer procedure) rather than the task that involved less pain (Kahneman et al., 1993). In that study and in ours, participants seemed to be using something other than memory for pain to make decisions for the future.

If people do not rely on their memory for how painful an experience was to decide how willing they are to repeat it again in the future, what informs that decision? Results from exploratory correlational analyses demonstrated that willingness to repeat the overall study again was related to memory for emotional experiences during the cold pressor, including recalled distress, positive affect, and negative affect. As for the emotional experiences reported *during* the cold pressor, willingness was only related to experienced positive affect, but not to experienced negative affect or distress. This implies that intentions to repeat past aversive experiences are more strongly related to memory for emotional responses to the aversive experience and not necessarily the actual emotional response that occurred in the moment or remembered pain. This is consistent with studies that have shown that intentions for future behavior rely more strongly on how a similar past experience is remembered than on how it was experienced in the

moment (e.g., Wirtz, Kruger, Napa Scollon, & Diener, 2003). Due to the exploratory nature of these analyses and the involvement of multiple comparisons, however, these results should be interpreted cautiously, and additional research will be necessary to assess the reliability of these findings.

These findings are particularly important in the healthcare domain, where decisions to pursue medical care in the future are based on memory for past healthcare experiences. Patients may be more likely to schedule a routine colonoscopy, for example, if they remember their previous colonoscopy as less aversive (Redelmeier et al., 2003). In the present study we aimed to examine whether changing memory for the pain experienced during an aversive task might make participants more willing to repeat that experience again in the future. Although participants exposed to misinformation exhibited greater distortion in their memories for their pain, we found that this alteration did not influence willingness to repeat the task again in the future. Perhaps changing memory for *emotional* responses to the aversive experience, rather than changing memory for physical pain, would be more effective in increasing willingness to repeat the experience in the future. This suggests that memory for pain may be used differently from memory for affective reactions to pain when it comes to making decisions for future behavior, even though pain and affective reactions to pain are similarly susceptible to memory distortion.

Detection measure

Following past research (Johansson et al., 2005), we included two measures of detection: concurrent and retrospective. *Concurrent* detectors spontaneously reported that the pain ratings they were shown were not the pain ratings they had truly made. This requires the participants to take initiative, and is perhaps one reason why we observed such a low rate of concurrent detection (11%). By contrast, *retrospective* detectors reported finding something strange about the process of the computer reminding them of their pain ratings. However, they had already been asked in the debriefing whether they had found anything strange or surprising about the study, which might have biased their response rate. Indeed, we found that a small number of control participants were coded as retrospective detectors, illustrating that this measure may have a higher false-positive rate. On the other hand, nondetectors had higher levels of social desirability than did detectors in both measures, meaning that some participants in the nondetector groups may have actually detected the misinformation but did not demonstrate this due to their desire to respond favorably to the researcher. The true rate of detectors likely lies somewhere between the measures of concurrent detection and retrospective detection employed in this study. Future research might attempt to design measures of detection that are less sensitive to demand characteristics.

Participants who detected the misinformation concurrently reported more distress, more negative affect, and marginally more pain than participants who did not detect the misinformation concurrently. This suggests that people may be more vigilant to misinformation about their own pain ratings the more distressing, negative, or painful their actual experience was (see also Hall, Johansson, & Strandberg, 2012). This is logical, as pain tends to grab a person's attention (Eccleston & Crombez, 1999). The more pain one experiences, the more they are likely to attend to how they feel in that moment, and the more likely they are to notice discrepancies between misinformation and their actual report.

Social desirability

In addition to examining how social desirability related to the tendency to detect misinformation, we also examined how social desirability related to our other dependent variables of interest. Higher levels of social desirability were not related to pain and distress reported during the cold pressor, but were related to recalling the cold pressor more favorably during the second session. This speaks to the potential impact of demand effects on reports of pain. Among the misinformation condition, social desirability was also related to a greater decrease in reports of pain during the second session, suggesting that people who have higher levels of social desirability are more susceptible to misinformation.

Limitations

Our study was limited by the fact that we examined pain in a laboratory setting, which may produce different responses than in a real-world setting. Participants might have been less attentive to alterations in their self-reported ratings in the laboratory than in a setting where their health is directly implicated, and thus may not have been as vigilant to misinformation. Furthermore, because the lab setting does not hold any real-life health implications, participants might have been less willing to repeat the same experience again than if they knew that repeating the procedure would potentially benefit their health. Therefore, if another study were conducted in a medical setting, we might expect higher misinformation detection rates, but also a greater willingness to repeat aversive procedures. Future research would do well to test memory bias for real-life procedures.

Conclusion

Memory of how a person felt in the past informs what that person is willing to do in the future. Memory is susceptible to bias, however, both from natural processes and external influences. Therefore, understanding the ways in which memory for past experiences might be biased is important for

predicting future behavior. This is particularly consequential in the healthcare domain, where patients may make medical decisions based on their memory for how painful a past experience was. The present study revealed that people can be misled to believe they experienced less pain than they actually reported during a cold pressor, and that this misinformation can become incorporated into their memories for the experience. In this way, we were able to "add a better end" by decreasing the amount of pain recalled from a painful experience. Unexpectedly, underestimated pain ratings did not translate to a greater willingness to repeat study procedures in the future. Instead, the recalled emotional reactions to the cold pressor, such as recalled distress, negative affect, and positive affect, were more strongly related to willingness to participate in the entire study procedure again. Therefore, memory for physical pain, although it was shown to be malleable to misinformation, may not be as integral to future decision making as is memory for emotional responses following the pain.

References

Bernstein, D. M., & Loftus, E. F. (2009). How to tell if a particular memory is true or false. *Perspectives on Psychological Science*, *4*, 370–374. https://doi.org/10.1111/j.1745-6924.2009.01140.x

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (1995). "I hardly cried when I got my shot!": Influencing children's reports about a visit to their pediatrician. *Child Development*, *66*, 193–208. https://doi.org/10.1111/j.1467-8624.1995.tb00865.x

Cochran, K. J., Greenspan, R. L., Bogart, D. F., & Loftus, E. F. (2016). Memory blindness: Altered memory reports lead to distortion in eyewitness memory. *Memory & Cognition*, *44*, 717–726. https://doi.org/10.3758/s13421-016-0594-y

Crowne, D. P., & Marlow, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349–354. https://doi.org/10.1037/h0047358

Eccleston, C., & Crombez, G. (1999). Pain demands attention: A cognitive–affective model of the interruptive function of pain. *Psychological Bulletin*, *125*, 356–366. https://doi.org/10.1037/0033-2909.125.3.356

Goyal, T. M., Shimbo, D., Mostofsky, E., & Gerin, W. (2008). Cardiovascular stress reactivity. In L. J. Luecken & L. C. Gallo (Eds.), *Handbook of physiological research methods in health psychology* (pp. 133–158). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412976244.n7

Hall, L., Johansson, P., & Strandberg, T. (2012). Lifting the veil of morality: Choice blindness and attitude reversals on a self-transforming survey. *PLoS ONE*, *7*, e45457. https://doi.org/10.1371/journal.pone.0045457

Hall, L., Johansson, P., Tärning, B., Sikström, S., & Deutgen, T. (2010). Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition*, *117*, 54–61. https://doi.org/10.1016/j.cognition.2010.06.010

JASP Team. (2018). JASP (Version 0.9) (Computer software). Retrieved from https://jasp-stats.org/download/

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, *310*, 116–119. https://doi.org/10.1126/science.1111709

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*, 401–405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x

Kahneman, D., & Riis, J. (2005). Living, and thinking about it: Two perspectives on life. In F. A. Huppert, N. Baylis, & B. Keverne (Eds.), *The science of well-being* (pp. 285–304). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198567523.003.0011

Levine, L. J., Lench, H. C., & Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Applied Cognitive Psychology*, *23*, 1059–1075. https://doi.org/10.1002/acp

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory*, *12*, 361–366. https://doi.org/10.1101/lm.94705

McFarland, C., Ross, M., & DeCourville, N. (1989). Women's theories of menstruation and biases in recall of menstrual symptoms. *Journal of Personality and Social Psychology*, *57*, 522–531. https://doi.org/10.1037/0022-3514.57.3.522

McLaughlin, O., & Somerville, J. (2013). Choice blindness in financial decision making. *Judgment and Decision Making*, *8*, 561–572. https://doi.org/10.1017/CBO9781107415324.004

Merckelbach, H., Dalsklev, M., Van Helvoort, D., Boskovic, I., & Otgaar, H. (2018). Symptom self-reports are susceptible to misinformation. *Psychology of Consciousness: Theory, Research, and Practice*, *5*, 384–397. https://doi.org/10.1037/cns0000159

Merckelbach, H., Jelicic, M., & Pieters, M. (2011). Misinformation increases symptom reporting: A test–retest study. *JRSM Short Reports*, *2*, 75. https://doi.org/10.1258/shorts.2011.011062

Mitchell, L. A., MacDonald, R. A. R., & Brodie, E. E. (2004). Temperature and the cold pressor test. *Journal of Pain*, *5*, 233–238. https://doi.org/10.1016/j.jpain.2004.03.004

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: *Pain, 66*, 3–8.

Redelmeier, D. A., Katz, J., & Kahneman, D. (2003). Memories of colonoscopy: a randomized trial. *Pain, 104*, 187–194. https://doi.org/10.1016/S0304-3959(03)00003-4

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934–960. https://doi.org/10.1037//0033-2909.128.6.934

Sauerland, M., Schell, J. M., Collaris, J., Reimer, N. K., Schneider, M., & Merckelbach, H. (2013). "Yes, I have sometimes stolen bikes": Blindness for norm-violating behaviors and implications for suspect interrogations. *Behavioral Sciences & the Law, 31*, 239–255. https://doi.org/10.1002/bsl.2063

Stille, L., Norin, E., & Sikström, S. (2017). Self-delivered misinformation—Merging the choice blindness and misinformation effect paradigms. *PLoS ONE, 12*, e173606:1–17. https://doi.org/10.1371/journal.pone.0173606

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14*, 329–338. https://doi.org/10.3758/BF03202511

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018a). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*, 58–76. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2018b). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review, 25*, 35–57. https://doi.org/10.3758/s13423-017-1343-3

Wirtz, D., Kruger, J., Napa Scollon, C., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science, 14*, 520–524. https://doi.org/10.1111/1467-9280.03455