Journal:

Behaviour Research And Therapy

Acceptance date:

21 November 2020

Doi:

<https://doi.org/10.1016/j.brat.2020.103777>

Title:

Differences in perceptual memory determine generalization patterns

Ann-Kathrin Zenses^a, Jessica C. Lee^b, Valérie Plaisance^a and Jonas Zaman^{a,c}

a Center for the Psychology of Learning and Experimental Psychopathology, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, Box 3712, 3000 Leuven, Belgium

b University of New South Wales, Sydney, NSW 2052, Australia

c Health Psychology, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, box 3726, 3000 Leuven, Belgium

Number of figures: 5

Number of Tables: 0

Word count: 5642

Supplemental information: yes

Corresponding author:

Jonas Zaman, PhD

Health psychology

Tiensestraat 102 - bus 3726

3000 Leuven

Tel. +32 16 37 31 97

Fax +32 16 3 26144

E-mail: Jonas.zaman@kuleuven.be

Institutional URL : <https://ppw.kuleuven.be/ogp>

Abstract

Although memory of past experiences is crucial for the ability to transfer knowledge to new situations, surprisingly little research has directly investigated the relationship between memory and generalization. The present study sought to investigate how the perceptual memory of a trained stimulus influences generalization to similar stimuli. Forty participants underwent a fear conditioning procedure on Day 1, and separate memory recall and generalization tests on Day 2. We focused on two aspects of perceptual memory: namely memory bias (i.e., over- or underestimation of stimulus magnitude) and uncertainty. We found that memory bias predicted the pattern of generalized self-reported (expectancy ratings) and psychophysiological responses (fear-potentiated startle responses). Memory uncertainty was measured in two ways: self-reported uncertainty ratings and variability in stimulus recall. We found that higher levels of self-reported memory uncertainty corresponded with a broader generalization gradient on US expectancy, while greater variability in memory recall was associated with a broader generalization gradient on fear-potentiated startle responses. Taken together, our findings suggest that memory is an important determinant of generalized behavior and illustrate the need to account for these interindividual differences in perceptual memory when examining the generalization of learned responses.

Keywords: Memory; Recall; Fear conditioning; Generalization; Uncertainty

1. Intro

One of the indispensable tools that we have at our disposal is the ability to use past experiences to guide current thoughts and actions in novel situations (called generalization). Numerous studies have demonstrated that the average strength (or probability) of a previously trained response wanes as a function of the physical difference between test and training stimuli, yielding a response gradient (Ghirlanda & Enquist, 2003; Mednick & Freedman, 1960).

Given that we do not have direct access to the physical properties of the stimuli, but must rely on the input from senses, theories of generalization include some mental representations of the stimuli from our environment (Blough, 1975; Ghirlanda & Enquist, 2007; McLaren & Mackintosh, 2002, Jäkel, Schölkopf, & Wichmann, 2007, 2008; Shepard, 1987; Shepard, 1958). A common assumption in these theories is that the memory representation of the trained stimulus is veridical. However, it is possible that there are individual differences in the accuracy of the memory representation. The degree of similarity, a crucial determinant of generalization, would therefore depend on how accurate participants could remember the trained stimulus (Zaman et al., in press). It follows that individual differences in perceptual memory, in addition to other processes, could play an important role in generalization. It has been even proposed that memory instability (i.e., the susceptibility of memory to interference and loss) is the pathway through which generalization originates (Robertson, 2018).

Stimulus memory may be disrupted in a number of ways, with differing effects on generalization. For instance, the forgetting of stimulus features due to the passage of time may lead to a widening of generalization gradients (Jasnow, Cullen, & Riccio, 2012; Riccio, Ackil, & Burch-Vernon, 1992; Riccio, Richardson, & Ebner, 1984). This proposal is based on animal research where longer delays between learning and testing flattened generalization gradients, while overall response strength did not fade (Perkins & Weyant, 1958; Thomas & Burr, 1969; Riccio et al., 1984). In those studies, memory effects were inferred from the widening of generalization gradients of a trained response but without any separate assessment of memory. Other studies have demonstrated that the variability of a memory representation increases with longer retention intervals (Huang & Sekuler, 2010). Yet, as none of these studies assessed both memory and generalization in separate tasks, their precise relationship remains to be empirically demonstrated.

One way to quantify individual differences in perceptual memory is to ask participants to recall certain physical characteristics of a stimulus during an adjustment task. In this task, subjects have to adjust the features of a stimulus so that it matches their representation in memory over multiple trials. Deviations from the actual stimulus tend to follow a normal distribution, with the mean and standard deviation describing the bias and variability of memory, respectively (Huang & Sekuler, 2010). If we

conceive of a participants' perceptual memory as a probability distribution, we may derive two predictions about how generalization will be affected (Riccio et al., 1984; Zaman et al., (in press)).

Deviations of the distribution's mean from the actual stimulus, reflective of a memory bias, could cause *asymmetries* in the response gradient (Zaman et al., in press). Indirect support for this comes from a study by Dunsmoor et al. (2009). They found that the test stimulus that elicited the strongest fear response was also most often identified as the trained stimulus during a recognition task (Dunsmoor, Mitroff, & LaBar, 2009). In addition, two recent studies combined a two-alternative forced-choice (2AFC) categorization task with a generalization task using different sized circles as stimulus set (Struyf, Zaman, Hermans, & Vervliet, 2017; in Supplementary Information [SI] of Zaman, Struyf, Ceulemans, Beckers, & Vervliet, 2019). The authors found that shifts in the probability of incorrectly categorizing test stimuli as the trained stimulus (potentially reflecting a memory bias) were associated with parallel shifts in the generalization gradient. In other words, a higher tendency to incorrectly categorize stimuli that were smaller or larger than the trained stimulus corresponded to higher generalized responding to those stimuli. Note that in these studies, generalization was tested immediately after training. With a longer delay between training and test, the influence of these memory biases on generalization may be even stronger.

The second prediction we may derive is that the width of the perceptual memory distribution, reflecting uncertainty, could increase the *breadth* of generalization (Jasnow et al., 2012; Riccio et al., 1992, 1984, Zaman et al., in press). A wider distribution should lead to more generalization because the overlap in probability distributions between the perceived stimulus features and recalled stimulus features would be larger than with a narrower distribution (Blough, 1975; Ghirlanda & Enquist, 2007). Furthermore, we might expect memory uncertainty to influence the effect of memory bias. That is, the more certain you are about your (biased) memory, the stronger its effect on generalized responding (Petzschner, Glasauer, & Stephan, 2015).

The present study aimed to determine the effect of perceptual memory bias and uncertainty on the generalization of an acquired fear response, across a range of differently sized circles (which have been used successfully to yield symmetrical generalization gradients; Zaman, Ceulemans, Hermans, & Beckers, 2019; Zaman, Struyf, et al., 2019). The experiment was spread over two days. We used a simple Pavlovian conditioning procedure with fear acquisition on the first day and assessment of stimulus memory (via recall) and generalization on the second day. In the acquisition phase, an image of a circle [conditioned stimulus (CS)] was repeatedly paired with a painful electrocutaneous stimulus [unconditioned stimulus (US)]. Twenty-four hours later, participants were asked to recall the circle size in an adjustment task and provide judgements of their memory uncertainty. Afterward, we

assessed generalization by measuring the extent to which a range of differently sized circles elicited the trained fear response (indicated by fear-potentiated startle responses and US expectancy ratings). Crucially, the recall phase occurred before the generalization test. That is, participants had to recall the size of the circle from their memory without being exposed to any stimuli, as would be the case if the recall phase had occurred after the generalization test or if we had instead used a recognition task to assess memory. We expected that both memory bias and uncertainty would affect the shape of the generalization gradient, with an asymmetrical gradient in the direction of the memory bias and a wider generalization gradient as uncertainty increased (Zaman et al., in press). Finally, we expected the effect of memory bias to disappear as uncertainty levels increased.

2. Method and Material

2.1. Preregistration

The study has been preregistered on the Open Science Framework (OSF) (<https://osf.io/qw4ke>). Deviations from the preregistered protocol and additional, exploratory analyses are reported as such. The complete data are publicly available on the OSF (<https://osf.io/42x9y/>).

2.2. Participants

Forty-one healthy volunteers participated in the study, of which 40 completed both parts and were included in the analyses [30 females (75%), mean age: 21.68 ($SD = 2.02$)]. The recruitment of participants took place through local advertisement boards and social media. As a compensation for their participation in both parts of the study, they could choose between 20 euros or partial course credits. Exclusion criteria, based on self-report, were cardiac or cardiovascular disorders, lung disease or breathing problems, neurological disorders, other severe medical conditions, (a history of) psychiatric disorders, acute or chronic pain, pregnancy, hearing difficulties, recommendation from a general practitioner to avoid stressful situations, any type of electronic implant, use of recreational drugs, recovering from severe trauma, taking medication that affects the central nervous system, being below eighteen or above fifty years of age and inadequate comprehension of the Dutch language. The study was approved by the social and societal ethics committee (SMEC) of the KU Leuven, Belgium (G-2016 10 641). All participants provided their written informed consent.

2.3. Electrocutaneous stimulation

Electrocutaneous stimuli were delivered through an electrocutaneous stimulation device (Constant Current Stimulator, model DS7), attached to the non-dominant wrist. The intensity of the stimulation was determined, for each individual, at the beginning of Day 1 and Day 2 by using the Ascending Method of Limits approach (Yarnitsky, Sprecher, Zaslansky, & Hemli, 1995). Participants were instructed that the stimulus should be painful but tolerable. Each stimulus was rated on a visual analog scale (VAS; 0 = *no pain*, 10 = *most imaginable pain*). Stimulation levels started at 1 mA and gradually increased until a VAS score of eight was obtained or the participant indicated that pain tolerance was reached. The average stimulation intensity was 20.48 mA ($SD = 8.85$) on Day 1 and increased to 24.65 mA ($SD = 11.94$) on Day 2 [$t(39) = 4.646$, $p < .001$].

2.4. Visual stimuli

Seven differently sized circles were created as white outlines against a black background (S1-S7, varying diameters from 7.72 to 12.56 cm with steps of 0.807 cm). Across a dimension of increasing diameter, the middle circle served as CS (diameter = 10.14 cm; S4), the others as generalization stimuli. Two rectangles (8.79 x 2.83 cm and 2.93 x 0.94 mm) were used as stimuli during the practice trials.

2.5. Eyeblink startle responses

Orbicularis oculi electromyographic (EMG) activity was recorded as index of covert defensive mobilization (Bradley & Lang, 2000; Lang, Davis, & Ohman, 2000) by using three trimmed disposable Ag/AgCl electrodes (H124SG, 24 mm, Covidien®, UK). Two of them were attached under the left eye and one on the forehead, according to the guidelines of Blumenthal and colleagues (Blumenthal et al., 2005). Before electrode attachment, the skin was abraded with a mild abrasive cream (Inecto). Startle eyeblink responding was triggered by a 50-ms burst of white noise (called the startle probe) with a peak of 105 dB presented binaurally through headphones (Sennheiser, HD418). The raw signal was passed through a 13Hz high pass and 500Hz low pass bandpass filter (v75-04, Coulbourn isolated bioamplifier®, USA), rectified and smoothed with a time constant of 20 ms (V76-24, Coulbourn Integrator®, USA). EMG was sampled at 1000 Hz and recorded continuously. Time intervals of 1000 ms prior to probe onset until 1000 ms after probe onset were used to visually inspect (offline) the EMG signal for artifacts and non-responses. Rejection criteria were spontaneous blinks in the 200 ms interval preceding probe onset or excessive noise in the EMG signal. First, startle eyeblink amplitudes were computed by subtracting the mean baseline value (0-20 ms after probe onset) from the peak value found in the 21-175 ms time window after startle probe onset. Next, to reduce inter-individual variation, startle amplitudes were transformed into *t*-scores. Finally, based on these *t*-scores, the difference between the startle response on a given trial from the averaged startle responses of noise alone trials (in total 12 trials) was computed to express responses relative to baseline (noise alone

trials). Due to excessive noise, 14.25% of the startle recording were rejected (342 of 2400 trials). If more than 50% of the startle trials were non-responses for a given participant, this participant was deemed a non-responder and their data omitted from the startle analyses ($n = 1$).

2.6. Protocol

2.6.1. Day 1

Participants were invited to attend two separate laboratory sessions over a two-day protocol with exactly 24 hours between sessions. On Day 1, participants gave informed consent after receiving written and oral information. The first day comprised a calibration, practice, startle probe habituation, habituation and acquisition phase (Fig. 1A). After determining the intensity of the US, participants were familiarized with the trial timings and the task over two practice trials. Each trial commenced with the presentation of a visual stimulus (i.e., one of the rectangles) simultaneously with the size-VAS (at the bottom of the screen; range: 0 - 200 with mm as unit, and a step size of 1) on which participants were asked to indicate the estimated diameter of the presented stimulus (Fig. 1B)¹. No feedback was provided about the accuracy of their response. After 5 s, the size-VAS disappeared and was replaced with the expectancy-VAS (at the bottom of the screen) for another 5 s during which the visual stimulus remained present. Participants were instructed to indicate their expectancy of the US on this scale using a mouse click (0 = *never*, 10 = *always*).

Fig. 1. (A) Overview of the different phases on Day 1. (B) Structure of a trial during the acquisition phase on Day 1. (C) Overview of the different phases on Day 2. (D) Structure of a trial during the recall

¹ This measure was included for exploratory analyses (see SI).

phase on Day 2. (E) Structure of a trial during the generalization phase on Day 2. The trial structure was identical to (B) with the only exception that generalization stimulus presentations could also occur instead of CS presentations. CS = conditioned stimulus.

Next, in the startle habituation phase, ten startle probes were presented with a variable interval between probes (average duration 20 s) while a fixation cross was presented on screen. The subsequent habituation phase consisted of six trials during which the CS was presented simultaneously with the size-VAS on which participants indicated the estimated diameter of the presented circle (Fig. 1B). After 5 s, the size-VAS disappeared and was replaced with the expectancy-VAS. Ten seconds after CS onset both the CS and the expectancy-VAS disappeared. Trials were separated by the presentation of a fixation cross with a variable duration between 8 and 12 s. There were no startle probe and US presentations during the habituation phase.

The acquisition phase started immediately after the habituation phase, and consisted of 20 CS trials and 2 noise alone trials. The trial structure was identical to the training phase with the only exception that the US and/or a startle probe could occur on certain trials (Fig. 1B). Thus, each trial commenced with the presentation of the CS and the size-VAS. After 5s, the size-VAS was replaced by the expectancy-VAS, which remained visible for another 5s until CS offset. A startle probe was presented on 40% of the trials during CS presentation (random between 5 and 9 s after CS onset). On 80% of the trials the CS was followed by the US, which was administered at CS and expectancy-VAS offset. During noise alone trials, the fixation cross remained on the screen and a startle probe was presented randomly between 4 and 8 s after fixation cross onset. Finally, participants completed the State-Trait Anxiety Inventory (STAI-T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) and the Groningen Sleep Quality Scale (Meijman, T. F., de Vries-Griever, De Vries, & Kampman, 1988) to measure trait anxiety and sleep quality during the previous night, respectively.²

2.6.2. Day 2

Day 2 consisted of a calibration, recall, startle probe habituation and generalization phase (Fig. 1C). The calibration procedure was identical on both days. The recall phase included 10 trials in which participants were instructed to recreate the size of the circle that was presented on Day 1 (i.e., the CS). They could do so by pressing the left and right key on the keyboard to decrease and increase the

² Due to practical reasons, the Pittsburgh Sleep Quality Index (PSQI; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989) could not be administered, although this was included in the preregistration.

diameter of a circle (in steps of 0.2929 mm) that started as a small dot, located at the center of the screen (Fig. 1D). Participants confirmed the recreated circle with a button press. While the recreated circle remained on screen, they indicated how certain they were about the correctness of their response (0 = *not certain*, 10 = *very certain*). Finally, they reported retrospective US expectancies (i.e., the likelihood that this circle was followed by the US during the acquisition phase) on a scale ranging from 0 (*never*) to 10 (*always*). Between trials, a fixation cross was presented with a variable duration between 8 and 12 s. At the end of this phase, participants were asked to indicate how many different circles they saw on Day 1. The generalization phase consisted of 5 blocks, separated by 1-min breaks, each block including 5 CS trials, 2 noise alone trials and 12 GS trials (two per GS type). Stimulus order within each block was random. Trial structure (including the reinforcement rate of 80% for CS trials) was identical to the acquisition phase on Day 1 (Fig. 1E). GS trials were identical to CS trials with the only exception that the GS was never followed by the US. A startle probe was presented in 47% of all trials, occurring randomly between 5 and 9 s after stimulus onset.

Before participants were debriefed, they completed the Groningen Sleep Quality Scale (Meijman, T. F. et al., 1988), which assessed participants' sleep quality of the previous night.

2.7. Data analyses

All analyses were performed using SPSS 20. Learning on Day 1 was assessed using mixed effects models. For the US expectancy ratings, an increase across trials was tested (Fixed effect: intercept + trial, Random effect: subject-dependent intercept). Startle eyeblink responses were expressed as difference scores from baseline responding (i.e., noise alone trials) and were tested across all trials (Fixed effect: intercept, Random effect: subject-dependent intercept). Startle eyeblink difference scores > 0 are indicative of startle potentiation (i.e., stronger startle responding during the stimulus trials than during the noise alone trials).

On Day 2, the data from the recall phase and generalization phase were analyzed separately. Potential changes across recall trials were assessed for all three measures: recalled CS size, recall uncertainty and retrospective US expectancy ratings (Fixed effect: intercept + trial, Random effect: subject-dependent intercept). For the generalization data (US expectancy ratings and startle eyeblink responses), we first assessed whether US expectancy rating differed across stimuli, yielding a response gradient (Fixed effect: intercept + stimulus + trial, Random effect: subject-dependent intercept).

Next, the influence of memory bias and uncertainty on the pattern of US expectancy ratings across the stimulus spectrum was assessed in separate models. Memory bias was operationalized as participants' average recalled CS size (μ_r) minus the actual CS size (memory bias = $\mu_r - 101.41$, see Fig.

3C). We tested the following model with fixed effects: intercept + trial + stimulus + memory bias + memory bias $\times$ stimulus and as random effect a subject-dependent intercept³. Two indices for memory uncertainty were computed: mean self-reported recall uncertainty (μ_u , Fig. 3D), and a more implicit index, the amount of variation in recalled size across a participant's recall trials (standard deviation, σ_r). Effects of memory uncertainty were tested in the following models, once with μ_u and once with σ_r (Fixed effect: intercept + trial + stimulus + memory uncertainty + memory uncertainty $\times$ stimulus, with the value of memory uncertainty being either μ_u or σ_r , Random effect: subject-dependent intercept). Finally, we investigated whether the influence of memory bias was dependent upon memory uncertainty. This was tested in two models depending on the memory uncertainty index (μ_u or σ_r) Fixed effect: intercept + trial + stimulus + memory uncertainty + memory bias + memory uncertainty $\times$ stimulus + memory bias $\times$ stimulus + memory bias $\times$ memory uncertainty $\times$ stimulus, with the value of memory uncertainty being either μ_u or σ_r , Random effect: subject-dependent intercept). Degrees of freedom were estimated using the Satterthwaite approximation (Satterthwaite, 1946) and were rounded off to the closest integer. Memory bias and memory uncertainty (μ_u or σ_r) were entered into all models as continuous predictors while stimulus was a categorical predictor⁴. Note that in contrast to the treatment of both predictors as continuous in the models, for visualization purposes and additional exploratory analyses we split the sample into groups. For memory bias, this was done by creating a negative and positive bias group, based on whether participants recalled the CS as smaller or larger. For memory uncertainty, we divided the sample into two groups using a median split. Finally, to visualize the interactions between memory bias and uncertainty, we used the models' predicted US expectancies based on the variable's mean $\pm$ 1 SD. We report the primary analyses relevant to our hypotheses. Parameter estimates and 95% CI intervals of the mixed models as well as other preregistered analyses can be found in the SI.

3. Results

3.1. Day 1

3.1.1. US expectancy ratings

³ For ease of interpretation, a difference score between mean recalled CS size and physical CS size was adopted for memory bias rather than what was preregistered (i.e., mean recalled CS size).

⁴ Deviating from the preregistration, a model with stimulus as a categorical predictor was adopted instead of a second-order polynomial with stimulus as continuous predictor, as this formalization was preferred based on the Bayesian Information Criterion (BIC) for both the US expectancy data (BIC_{cat}: 15132 vs. BIC_{pol}: 15149) and the startle data (BIC_{cat}: 9694 vs. BIC_{pol}: 9709), with lower BIC values indicating better fit. The parameter estimates from the original analyses can be found in the SI.

The pairing of the CS with the US led to an increase in US expectancy ratings across acquisition trials [trial effect: $F(1,771) = 315.84, p < .001$] (Fig. 2A).

3.1.2. Eyeblink startle responses

Across the acquisition phase, startle responses were potentiated during the CS compared to baseline responding [intercept: $F(1,39) = 24.699, p < .001$] (Fig. 2B).

Fig. 2. (A) Averaged US expectancy ratings across acquisition trials (red line). The grey lines are US expectancy ratings per participant. **(B)** Averaged startle eyeblink responses (red line) expressed as a difference score from baseline responding (i.e., average during noise alone trials (ITI) across both days). The grey squares are the averaged startle eyeblink response per participant. Error bars denote standard errors. US = unconditioned stimulus, ITI = inter-trial interval

3.2. Day 2

3.2.1. Recall phase

3.2.1.1. Recalled CS size

On the first trial of the recall phase, the averaged recalled CS size was 84.54 mm ($SD = 25.64$) and increased to 96.91 mm ($SD = 15.17$) on the last trial [trial effect: $F(1,358) = 23.574, p < .001$]. Figure 3A shows that, recall accuracy increased slightly over trials and approached the veridical CS size of 101.41 mm.

3.2.1.2. Recall uncertainty

There was a significant main effect of trial [trial effect: $F(1,359) = 48.11, p < .001$] with ratings become more certain across trials (Fig. 3B).

3.2.1.3. Retrospective US expectancy ratings

The average retrospective US expectancy rating was 6.40 ($SD = 2.22$; Fig 3E). Retrospective US expectancy did not significantly change across recall trials [trial effect: $F(1,359) = 0.028$, $p = .867$].

Fig. 3. (A-B) Recalled CS size and recall uncertainty ratings across trials averaged across participants (red) and per participant (grey). **(C-E)** The red line indicates the mean memory bias [calculated as the difference between mean recalled size (μ_r) and the CS diameter], mean self-reported recall uncertainty (μ_u) and mean retrospective US expectancy ratings during the recall phase on Day 2. The grey line in (C) denotes the physical size of the CS. The black squares are mean scores over trials for each participant. Error bars denote standard errors. CS = conditioned stimulus, US = unconditioned stimulus.

3.2.2. Generalization phase

3.2.2.1. US expectancy ratings

We found significant differences in US expectancy ratings between test stimuli were found [stimulus effect $F(6, 3362) = 38.483, p < .001$] (Fig. 4A). In exploratory analyses, we tested whether the gradient was bell-shaped, and found a significant quadratic trend [$t(3363) = -13.06, p < .001$]. US expectancy ratings decreased across trials [$F(1,3362) = 55.967, p < .001, \beta = -0.01, 95\% CI = (-0.01, -0.01)$]. Next, we investigated separately the influence of memory bias and uncertainty (μ_u or σ_r) through their inclusion and interaction with stimulus.

Memory bias did not influence overall US expectancy ratings [memory bias effect: $F(1,38) = 0.452, p = .505$] but did affect ratings differently across the stimulus range [memory bias $\times$ stimulus effect: $F(6,3356) = 2.594, p = .016$]. In exploratory follow-up analyses, we investigated the asymmetry in the gradient for subjects that recalled the CS as smaller (negative memory bias, $n = 33$) or as larger (positive memory bias, $n = 7$). A positive memory bias led to an asymmetrical gradient as the right-hand side of the gradient was higher than the left-hand side of the gradient [left vs. right side: $t(2365) = 2.382, p = .017$]. For the negative memory bias, no evidence for such an asymmetrical gradient was observed [left vs. right side: $t(2365) = 1.023, p = .305$] (Fig. 4B).

The effect of memory uncertainty was assessed using two measures: the individual's variation in recalled CS size (σ_r) and the averaged recall uncertainty ratings (μ_u). There was no significant effect of variation in recalled CS size (σ_r) [recall variation effect: $F(1,38) = 0.167, p = .685$; recall variation $\times$ stimulus effect: $F(6,3355) = 1.00, p = .422$]. Self-reported recall uncertainty (μ_u) did influence US expectancy ratings across the stimulus range but had no effect on the overall US expectancy ratings [recall uncertainty $\times$ stimulus effect: $F(6,3356) = 5.116, p < .001$; recall uncertainty effect: $F(1,38) = 0.835, p = .367$]. In exploratory follow-up analyses, we split the sample into a *high* and *low* recall uncertainty group using a median split (median = 6.4). As can be seen in Fig. 4C, the gradient flattens for *high* recall uncertainty compared to the *low* uncertainty, supported by a significant interaction with quadratic trend [Quadratic contrast vs. recall uncertainty group: $t(3357) = 4.961, p < .001$].

Finally, effects of memory bias and uncertainty (and their interaction) were tested simultaneously. Memory uncertainty assessed by self-report (μ_u) only influenced the interaction between memory bias and stimulus [recall uncertainty $\times$ stimulus $\times$ memory bias effect: $F(6,3344) = 2.306, p = .032$; recall uncertainty $\times$ memory bias effect: $F(1,36) < 1, p = .983$]. For illustrative purposes, we have plotted the model's predicted US expectancy gradients for relative *low* (uncertainty rating of 7.62, based on $\mu_u + 1 SD$ with $\mu_u = 6.06$ and $SD = 1.56$) and *high* levels of recall uncertainty (uncertainty of 4.50, based on $\mu_u - 1 SD$), and for a positive and negative memory bias set to 12.56 and -12.56 (12.56 is the average of the absolute values of the recall difference scores to reflect the overall strength of bias regardless of its direction, see Fig 4D). Although it appears that higher uncertainty was associated with flatter

gradients for both positive and negative memory bias, this effect of uncertainty was much more pronounced for positive than for negative bias (see Figure 4D). In contrast, the more implicit measure of recall variation did not influence the effect of memory bias on US expectancy responses in any manner [recall variation $\times$ stimulus $\times$ memory bias effect: $F(6,3344) = 1.424$, $p = .201$; recall variation $\times$ memory bias effect: $F(1,36) = 0.267$, $p = .608$].

Fig. 4. (A) Averaged US expectancy ratings per stimulus across participants (red) and per participant (grey) from the generalization phase on Day 2. **(B)** Averaged US expectancies for a positive memory bias (participants that recalled the CS as larger than its actual size, $n = 7$) and a negative memory bias

(participants that recalled the CS as smaller than its actual size, $n = 33$). **(C)** *Averaged US expectancy ratings for the low (≤ 6.4) and high self-reported recall uncertainty (μ_u) groups (> 6.4) based on a median split.* **(D)** *Predicted US expectancies of the mixed model based on the following values for memory bias and self-reported recall uncertainty : a positive (12.56, based on the absolute mean memory bias) and negative memory bias (-12.56) and low and high levels of self-reported recall uncertainty (based on a mean ± 1 SD with mean = 6.06 and SD = 1.56). Error bars denote standard errors. S = generalization stimuli, CS = conditioned stimulus, US = unconditioned stimulus, pos. = positive, neg. = negative.*

3.2.2.1. Startle eyeblink responses

Startle responses did not differ across the tested stimulus dimension [stimulus effect: $F(6,1274) = 0.387$, $p = .888$], although responses were potentiated compared to the baseline [intercept: $F(1,98) = 50.161$, $p < .001$]. Responses habituated across trials [trial effect: $F(1,1274) = 64.314$, $p < .001$] (Fig. 5A).

Next, effects of memory bias and uncertainty (and their interaction) were assessed. Memory bias did not influence the strength of startle responding [memory bias effect: $F(1,37) = 0.396$, $p = .533$] but did affect response strength differently across the stimulus range [memory bias $\times$ stimulus effect: $F(6,1268) = 2.581$, $p = .017$]. This effect was further explored in a similar way to the US expectancy analyses (Fig. 5B). However, these explorative analyses yielded no evidence for an asymmetry in either the positive or the negative bias group [pos. bias (left vs. right side): $t(242) = .624$, $p = .534$; neg. bias (left vs. right side): $t(1026) = .337$, $p = .891$]. Additional explorative tests yielded no difference in responding between any of the stimuli in the negative memory group (all p 's = ns). In the positive memory group, responding differed only between S1 and the CS $t(242) = 3.014$, $p = .003$, all other tests were non-significant after Bonferonni correction (all p 's = ns).

Memory uncertainty assessed via self-report or variation in recalled size, did not affect responding in general [memory uncertainty : $F(1,38) = 0.422$, $p = .520$; memory variation: $F(1,37) = 0.776$, $p = .384$] nor did it interact with stimulus [memory uncertainty $\times$ stimulus effect: $F(6,1269) = 0.578$, $p = .748$; memory variation $\times$ stimulus effect: $F(6,1268) = 0.910$, $p = .487$].

Lastly, we tested for effects of memory bias and uncertainty and their interaction. Memory uncertainty assessed by self-report did not influence the effect of memory bias on response strength across the stimulus range [memory uncertainty $\times$ memory bias effect: $F(1,35) = 3.217$, $p = .081$; memory uncertainty $\times$ memory bias $\times$ stimulus effect: $F(6,1256) = 1.212$, $p = .972$], however, there was a significant three-way interaction between memory variation, memory bias and stimulus [memory

variation $\times$ memory bias $\times$ stimulus effect: $F(6,1256) = 2.979$, $p = .007$; memory variation $\times$ memory bias effect: $F(1,36) = 0.385$, $p = .539$]. Although the gradients are quite variable in the positive recall bias panel, it can be seen (Fig. 5C) that patterns of startle responding became more similar between both types of memory bias when memory variation was high, whereas with a low memory variation pattern startle responses differed between positive and negative memory bias.

Fig. 5. (A) Startle eyeblink responses per stimulus across participants (red) and per participant (grey) from the generalization phase on Day 2. Startle responses are expressed as difference scores from the mean startle responding during noise alone trials (ΔITI , inter-trial interval). **(B)** Averaged startle eyeblink responses for a positive memory bias (recalling the CS as larger) and a negative memory bias (recalling the CS as smaller). **(C)** Predicted startle eyeblink responses of the mixed model for the combination between positive (12.56) and negative memory bias (-12.56) and low and high levels of

recall variation (based on $\mu_\sigma \pm 1$ SD with $\mu_\sigma = 8.30$ and $SD = 9.14$). Error bars denote standard errors. S = generalization stimuli, CS = conditioned stimulus, pos. = positive, neg. = negative, ITI = inter-trial interval

4. Discussion

This study sought to investigate the role of perceptual memory bias and uncertainty on the generalization of a trained fear response across a range of physical similar stimuli. It differs from the typical generalization study in separating training and test phases by 24 hours and presenting a recall test of the training stimulus prior to assessing generalization. To the best of our knowledge, our experiment is the first to provide direct evidence that uncertainty of the perceptual memory of a trained stimulus results in broader response generalization. Further, we found some evidence that memory bias (over- or under-recalling of the stimulus size) influenced the pattern of generalized self-reported (US expectancy ratings) and psychophysiological responding (fear potentiated startle responses). Our findings suggest that individual differences in stimulus memory are an important determinant of generalized behavior.

Our first set of novel findings concern the effect of memory bias. In our study, memory bias was assessed as the mean deviation during a recall adjustment task similar to previous studies (Allred, Crawford, Duffy, & Smith, 2016; Huang & Sekuler, 2010). Participants that recalled the CS as being larger than it was (a positive memory bias) also showed higher generalised responses to larger stimuli. In other words, generalized responding is determined by the degree to which a test stimulus is similar to participants' representation of the CS, which may be biased. Memory bias may thus be one explanation for an asymmetrical gradient obtained after conditioning with a single CS. This pattern of results is consistent with two previous studies that used a 2AFC categorization task (Struyf, Zaman, Hermans, & Vervliet, 2017; in SI of Zaman, Struyf, Ceulemans, Beckers, & Vervliet, 2019). However, unlike previous studies that also found an effect of a negative memory bias, we did not find higher generalised responses to smaller stimuli for participants who recalled the CS as being smaller (a negative memory bias). One potential reason for the observed differential effect of memory bias is that the starting point (the anchor) in the adjustment task of the recall phase was always a small dot. LeBoef and Shafir (2006) demonstrated in a series of six studies that decreasing the size of an anchor, during an adjustment task, reliably led to size underestimations. This anchoring effect may explain why

such a large proportion of participants in our study showed a negative memory bias (33/40) (Allred et al., 2016; LeBoeuf & Shafir, 2006). Future studies should therefore use randomized or counterbalanced starting points to avoid potential anchoring effects.

We also found an effect of memory bias on the startle potentiation gradients, although the results were less clear. A potential issue is that the extreme flat startle gradient might have acted as a ceiling effect. Upon testing on Day 2, all test stimuli elicited potentiated startle responses that did not differ from the initial trained fear response, indicating a high level of generalized fear. Hence, a potential asymmetry in response decline may be concealed by the absence of any response differentiation, indicative of maximal generalization. Although not designed to test for an effect of the delay between learning and testing, differences in gradients between this study and similar studies that tested generalization immediately after learning (Struyf, Zaman, Hermans, & Vervliet, 2017; Zaman, Struyf, Ceulemans, Beckers, & Vervliet, 2019) suggest that the 24h delay period in the current study contributed to the flattening of the gradient. In our study, both US expectancy and startle gradients were much flatter (while still remaining high) compared to previous studies (Struyf, Zaman, Hermans, & Vervliet, 2017; Zaman, Struyf, Ceulemans, Beckers, & Vervliet, 2019), suggestive of more generalization. This observation fits with the proposal that generalization gradients flatten over time due to increased memory variability (Thomas & Burr, 1969; Riccio et al., 1984).

Our second set of findings concern the effect of memory uncertainty. The key result was that memory uncertainty was related to flatter generalization gradients. This finding is in line with previous suggestions that the width of the memory distribution, reflective of its precision, has been proposed to increase the extent of generalization (Jasnow et al., 2012; Riccio et al., 1992, 1984). However, this theoretical account was predominantly based on indirect empirical evidence from animal research where longer delays between learning and testing flattened generalization gradients (Perkins & Weyant, 1958). Our experiment is the first to provide direct evidence that memory uncertainty results in more generalization behavior. We operationalized memory uncertainty in two ways: the amount of variation during the adjustment task (more implicit measure), and self-report (more explicit measure). Interestingly, we found differential effects of these indices. For US expectancy gradients, self-reported memory uncertainty broadened generalization, whereas for startle potentiation the effect was only found for recall variation (only in interaction with the memory bias effect). Interestingly, our two measures of memory uncertainty did not correlate (see SI), suggesting they relate to different components of uncertainty. Conceptually, obtaining memory uncertainty using self-report ratings closely relates to the concept of metamemory cognitions (i.e., awareness about one's own memory

capabilities) whereas response variation across recall trials does not (Nelson, 1990). Research should further investigate the extent to which these different aspects of memory uncertainty may explain previously reported divergences between explicit and implicit indices of fear learning (Holt et al., 2014; Sevenster, Beckers, & Kindt, 2012, 2014).

Although we demonstrated inter-individual differences in the memory of a fear evoking stimulus and its influence on fear generalization, the underlying causes of these individual differences and their theoretical and clinical relevance remain to be elucidated. In real life, determining the optimal extent of generalization requires subtle balance: too little generalization will result in suboptimal use of previous information, whereas excess generalization will result in cognitive and behavioral rigidity (Jäkel et al., 2008), which has been implicated in various pathologies including anxiety disorders (Kindt, 2014; Lissek, 2012). Insight into the determinants of this process is therefore of clinical importance as many different mechanisms including inaccurate perception, biased memory, or rigid cognitions may all yield the same seemingly pathological behavior, yet require distinct treatment forms.

Unlike previous studies, a key novel contribution of the current study was the assessment of memory and generalization using separate tasks. This enabled us to directly investigate the impact of memory on generalization. Furthermore, the use of a recall task instead of a recognition task enabled us to investigate perceptual memory in a detailed manner without presenting opportunities for further learning and extinction. However, it seems that the repeated probing of participant's memory had some effect on judgements despite the fact that no feedback was provided during the recall task. Across recall trials subjects became more accurate and also more certain. This may be explained by the anchoring effect discussed above. Participants' judgements may have been biased by the small anchor point and gradually became less and less susceptible to the anchor over the repeated trials. Further, although it was necessary to reinforce the CS during generalization testing to prevent extinction, we acknowledge that this may have provided participants with new learning opportunities to correct any potential memory bias. We note, however, that if anything, this should have worked *against* us finding any effect of memory bias on generalization. Future studies are required to determine the optimal design for concurrent assessment of memory and generalization.

In sum, we have demonstrated that differences in the memory of an identical experience influence the extent to which people generalize from it. Models of generalization that assume a veridical representation between stimuli and their memory representation may account for group-level data, but fail to account for important differences between individuals or subgroups of participants (Zaman et al., in press., see also Lee, Hayes, & Lovibond, 2018; Lovibond, Lee, & Hayes, 2020). Our data

illustrate the need to account for these interindividual differences in perceptual memory when examining generalization of learned responses.

Acknowledgments: JL received salary support from the Australian Research Council (DP190103738). JZ is a Postdoctoral Research Fellow of the Research Foundation Flanders (FWO, 12P8619N) and received funding from the Efic-Grunthal grant (EGG ID 358254826), a Special Research Funds (FWO, 1500620N) and the KU Leuven Research Council (POR/20/001). The research was sponsored by the “Asthenes” long-term structural funding (METH/15/011) – Methusalem grant by the Flemish Government.

References

- Allred, S. R., Crawford, L. E., Duffy, S., & Smith, J. (2016). Working memory and spatial judgments: Cognitive load increases the central tendency bias. *Psychonomic Bulletin & Review*, 23(6), 1825–1831. <https://doi.org/10.3758/s13423-016-1039-0>
- Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 1(1), 3–21.
- Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, 42(1), 1–15. <https://doi.org/10.1111/j.1469-8986.2005.00271.x>
- Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37(2), 204–215. <https://doi.org/10.1111/1469-8986.3720204>
- Buyse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
- Dunsmoor, J. E., Mitroff, S. R., & LaBar, K. S. (2009). Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning & Memory*, 16(7), 460–469. <https://doi.org/10.1101/lm.1431609>
- Dunsmoor, Joseph E., Mitroff, S. R., & LaBar, K. S. (2009). Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning and Memory*, 16(7), 460–469. <https://doi.org/10.1101/lm.1431609>
- Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, 66(1), 15–36. <https://doi.org/10.1006/anbe.2003.2174>
- Ghirlanda, S., & Enquist, M. (2007). How training and testing histories affect generalization: A test of simple neural networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1479), 449–454. <https://doi.org/10.1098/rstb.2006.1972>
- Holt, D. J., Boeke, E. A., Wolthusen, R. P. F., Nasr, S., Milad, M. R., & Tootell, R. B. H. H. (2014). A parametric study of fear generalization to faces and Non-Face objects: Relationship to discrimination thresholds. *Frontiers in Human Neuroscience*, 8(SEP), 1–12. <https://doi.org/10.3389/fnhum.2014.00624>
- Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: two classes of attractors at work. *Journal of Vision*, 10(2), 24.1-27. <https://doi.org/10.1167/10.2.24>
- Jäkel, F., Schölkopf, B., & Wichmann, F. A. (2007). A tutorial on kernel methods for categorization. *Journal of Mathematical Psychology*, 51(6), 343–358. <https://doi.org/10.1016/j.jmp.2007.06.002>
- Jäkel, F., Schölkopf, B., & Wichmann, F. A. (2008). Generalization and similarity in exemplar models of categorization: Insights from machine learning. *Psychonomic Bulletin and Review*, 15(2), 256–271. <https://doi.org/10.3758/PBR.15.2.256>
- Jasnow, A. M., Cullen, P. K., & Riccio, D. C. (2012). Remembering another aspect of forgetting. *Frontiers in Psychology*, 3(JUN), 1–8. <https://doi.org/10.3389/fpsyg.2012.00175>
- Kindt, M. (2014). A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behaviour Research and Therapy*, 62, 24–36. <https://doi.org/10.1016/j.brat.2014.08.012>

- Lang, P. J., Davis, M., & Ohman, A. (2000). Fear and anxiety: animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, 61(3), 137–159. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11163418>
- LeBoeuf, R. A., & Shafir, E. (2006). The long and short of it: physical anchoring effects. *Journal of Behavioral Decision Making*, 19(4), 393–406. <https://doi.org/10.1002/bdm.535>
- Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018). Peak shift and rules in human generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 1955–1970. <http://dx.doi.org/10.1037/xlm0000558>
- Lissek, S. (2012). Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of Pavlovian fear-learning: the case for conditioned overgeneralization. *Depression and Anxiety*, 29(4), 257–263. <https://doi.org/10.1002/da.21922>
- Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of classically conditioned-fear. *Biological Psychiatry*, 75(11), 909–915. <https://doi.org/10.1016/j.biopsych.2013.07.025>
- Lovibond, P. F., Lee, J. C., & Hayes, B. K. (2020). Stimulus discriminability and induction as independent components of generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46, 1106–1120. <https://doi.org/10.1037/xlm0000779>
- McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, 30(3), 177–200. <https://doi.org/10.3758/BF03192828>
- Mednick, S. A., & Freedman, J. L. (1960). Stimulus generalization. *Psychological Bulletin*, 57(3), 169–200. <https://doi.org/10.1037/h0041650>
- Meijman, T. F., de Vries-Griever, A. H., De Vries, G., & Kampman, R. (1988). The evaluation of the Groningen sleep quality scale. *Groningen: Heymans Bulletin (HB 88-13-EX)*, 2006.
- Nelson, T. O. (1990). Metamemory: A Theoretical Framework and New Findings. *Psychology of Learning and Motivation - Advances in Research and Theory*, 26(C), 125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Perkins, C. C., & Weyant, R. G. (1958). The interval between training and test trials as a determiner of the slope of generalization gradients. *Journal of Comparative and Physiological Psychology*, 51(5), 596–600. <https://doi.org/10.1037/h0042550>
- Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, 19(5), 285–293. <https://doi.org/10.1016/j.tics.2015.03.002>
- Riccio, D. C., Ackil, J., & Burch-Vernon, A. (1992). Forgetting of stimulus attributes: methodological implications for assessing associative phenomena. *Psychological Bulletin*, 112(3), 433–445. <https://doi.org/10.1037/0033-2909.112.3.433>
- Riccio, D. C., Richardson, R., & Ebner, D. L. (1984). Memory retrieval deficits based upon altered contextual cues: A paradox. *Psychological Bulletin*, 96(1), 152–165. <https://doi.org/10.1037/0033-2909.96.1.152>
- Robertson, E. M. (2018). Memory instability as a gateway to generalization. *PLoS Biology*, 16(3), 1–16. <https://doi.org/10.1371/journal.pbio.2004633>
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics*, 2(6), 110–114. <https://doi.org/10.2307/3002019>

- Sevenster, D., Beckers, T., & Kindt, M. (2012). Retrieval per se is not sufficient to trigger reconsolidation of human fear memory. *Neurobiology of Learning and Memory*, 97(3), 338–345. <https://doi.org/10.1016/j.nlm.2012.01.009>
- Sevenster, D., Beckers, T., & Kindt, M. (2014). Fear conditioning of SCR but not the startle reflex requires conscious discrimination of threat and safety. *Frontiers in Behavioral Neuroscience*, 8. <https://doi.org/10.3389/fnbeh.2014.00032>
- Shepard, R. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323. <https://doi.org/10.1126/science.3629243>
- Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, 65(4), 242–256. <https://doi.org/10.1037/h0043083>
- Spielberger, C. D., Gorsuch, R. C., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.
- Struyf, D., Zaman, J., Hermans, D., & Vervliet, B. (2017). Gradients of fear: How perception influences fear generalization. *Behaviour Research and Therapy*, 93, 116–122. <https://doi.org/10.1016/j.brat.2017.04.001>
- Thomas, D. R., & Burr, D. E. S. (1969). Stimulus generalization as a function of the delay between training and testing procedures: A reevaluation. *Journal of the Experimental Analysis of Behavior*, 12(1), 105–109. <https://doi.org/10.1901/jeab.1969.12-105>
- Yarnitsky, D., Sprecher, E., Zaslansky, R., & Hemli, J. a. (1995). Heat pain thresholds: normative data and repeatability. *Pain*, 60(3), 329–332. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7596629>
- Zaman, J., Anastasia, A., Zenses, A., Bilgim, A., Beckers, T., Vervliet, B., & Boddez, Y. (n.d.). Perceptual Variability: Implications for Learning and Generalization. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-020-01780-1>
- Zaman, J., Ceulemans, E., Hermans, D., & Beckers, T. (2019). Direct and indirect effects of perception on generalization gradients. *Behaviour Research and Therapy*, 114(January), 44–50. <https://doi.org/10.1016/j.brat.2019.01.006>
- Zaman, J., Struyf, D., Ceulemans, E., Beckers, T., & Vervliet, B. (2019). Probing the role of perception in fear generalization. *Scientific Reports*, 9(1), 10026. <https://doi.org/10.1038/s41598-019-46176-x>

Supplemental information

Day 1 - acquisition	US expectancy		Startle EMG		Size	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	4.97 (0.228)***	4.515, 5.428	6.338 (1.275)***	3.758, 8.920	68.394 (4.821)	58.655, 78.132
Trial	0.197 (0.011)***	0.175, 0.219	-	-	0.184 (0.064)	0.052, 0.310
Random effects						
Variance RI	1.427 (.356)	-	47.814 (14.427)	-	905.92 (206.368)	-
Residual variance	3.329 (.169)	-	103.074 (9.570)	-	104.153 (5.44)	-

RI = Random Intercepts, CI = Confidence Interval, ^ $\leq .1$, * $\leq .05$, ** $\leq .01$, *** $\leq .001$

Day 2 - recall	Size		Uncertainty		US expectancy	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	89.103 (2.570)***	83.961, 94.244	5.268 (0.275)***	4.719, 5.818	6.408 (0.329)	5.748, 7.068
Trial	0.997 (0.205)***	0.593, 1.400	0.144 (0.021)***	0.103, 0.185	0.184 (0.064)	-0.042, 0.036
Random effects						
Variance RI	199.109 (48.237)	-	2.350 (0.565)	-	3.71 (0.870)	-
Residual variance	138.143 (10.325)	-	1.421 (0.106)	-	1.30 (0.097)	-

RI = Random Intercepts , CI = Confidence Interval, ^ $\leq .1$, * $\leq .05$, ** $\leq .01$, *** $\leq .001$

Day 2 - US expectancy						
	Model 1		Model 2		Model 3	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	4.670 (0.223)***	4.226, 5.114	4.670 (0.224)***	4.224, 5.117	4.670 (0.224)***	4.224, 5.117
Trial	-0.011 (0.002)***	-0.014, -0.008	-0.011 (0.002)***	-0.015, -0.008	-0.011 (0.002)***	-0.015, -0.008
Stimulus (S1)	-0.469 (0.153)**	-0.769, -0.168	-0.469 (0.153)**	-0.769, -0.168	-0.469 (0.153)**	-0.769, -0.168
Stimulus (S2)	0.210 (0.153)	-0.090, 0.512	0.211 (0.153)	-0.090, 0.512	0.211 (0.153)	-0.090, 0.512
Stimulus (S3)	1.006 (0.153)***	0.841, 1.511	1.006 (0.153)***	0.706, 1.307	1.006 (0.153)***	0.705, 1.307
Stimulus (S4 = CS)	1.147 (0.128)***	0.896, 1.399	1.147 (0.128)***	0.896, 1.399	1.147 (0.128)***	0.896, 1.399
Stimulus (S5)	0.795 (0.153)***	0.494, 1.096	0.796 (0.153)***	0.495, 1.096	0.796 (0.153)***	0.494, 1.096
Stimulus (S6)	0.226 (0.153)	-0.075, 0.526	0.226 (0.153)	-0.074, 0.526	0.226 (0.153)	-0.075, 0.526
Stimulus (S7)	-	-	-	-	-	-
Bias			0.029 (0.015)^	-0.001, 0.058		
S1 × Bias			-0.036 (0.011)***	-0.057, -0.015		
S2 × Bias			-0.031 (0.011)**	-0.051, -0.010		
S3 × Bias			-0.013 (0.011)	-0.034, 0.007		
S4 × Bias			-0.024 (0.009)**	-0.041, -0.006		
S5 × Bias			-0.014 (0.011)	-0.034, 0.007		
S6 × Bias			-0.022 (0.011)*	-0.043, -0.002		
S7 × Bias			-	-		
SD _R					0.019 (0.024)	-0.029, 0.067
S1 × SD _R					-0.013 (0.017)	-0.046, 0.021
S2 × SD _R					0.005 (0.017)	-0.028, 0.039
S3 × SD _R					-0.023 (0.017)	-0.056, 0.011
S4 × SD _R					-0.023 (0.014)	-0.051, 0.005
S5 × SD _R					-0.009 (0.017)	-0.043, 0.024
S6 × SD _R					-0.011 (0.017)	-0.044, 0.022
S7 × SD _R					-	-
Random effects						
Variance RI	1.370 (0.327)	-	1.371 (0.327)	-	1.382 (0.330)	-
Residual variance	4.751 (0.116)	-	4.710 (0.115)	-	4.724 (0.115)	-

RI = Random Intercepts, CI = Confidence Interval, SD_R = standard deviation during recall, S7 serves as reference ^ ≤.1, * ≤.05, ** ≤.01, *** ≤.001

Day 2 - US expectancy						
	Model 4		Model 5		Model 6	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	4.671 (0.223)***	4.225, 5.117	4.672 (0.227)***	4.220, 5.125	4.672 (0.224)***	4.224, 5.117
Trial	-0.011 (0.002)***	-0.014, -0.008	-0.011 (0.002)***	-0.014, -0.008	-0.011 (0.002)***	-0.015, -0.008
Stimulus (S1)	-0.470 (0.153)**	-0.769, -0.170	-0.470 (0.153)**	-0.769, -0.171	-0.472 (0.153)**	-0.769, -0.168
Stimulus (S2)	0.210 (0.153)	-0.090, 0.509	0.210 (0.153)	-0.090, 0.508	0.199 (0.153)	-0.090, 0.512
Stimulus (S3)	1.004 (0.153)***	0.705, 1.304	1.003 (0.153)***	0.706, 1.302	0.990 (0.153)***	0.705, 1.307
Stimulus (S4 = CS)	1.146 (0.128)***	0.896, 1.397	1.145 (0.128)***	0.895, 1.395	1.138 (0.128)***	0.896, 1.399
Stimulus (S5)	0.794 (0.153)***	0.494, 1.096	0.793 (0.153)***	0.495, 1.096	0.784 (0.153)***	0.494, 1.096
Stimulus (S6)	0.224 (0.153)	-0.075, 0.523	0.223 (0.153)	-0.074, 0.526	0.224 (0.153)	-0.075, 0.526
Stimulus (S7)	-	-	-	-	-	-
Bias			0.026 (0.016)	-0.006, 0.058	0.028 (0.016)^	-0.005, 0.060
S1 × Bias			-0.037 (0.011)***	-0.059, -0.015	-0.035 (0.011)**	-0.057, -0.012
S2 × Bias			-0.031 (0.011)**	-0.052, -0.008	-0.024 (0.011)*	-0.047, -0.0002
S3 × Bias			-0.008 (0.011)	-0.030, 0.014	-0.005 (0.011)	-0.027, 0.018
S4 × Bias			-0.016 (0.009)^	-0.033, 0.003	-0.019 (0.009)*	-0.037, -0.001
S5 × Bias			-0.010 (0.011)	-0.032, 0.012	-0.008 (0.011)	-0.030, 0.014
S6 × Bias			-0.019 (0.011)^	-0.041, 0.002	-0.022 (0.011)^	-0.044, 0.001
S7 × Bias			-	-	-	-
SD _R					0.024 (0.026)	-0.028, 0.077
S1 × SD _R					-0.019 (0.019)	-0.055, 0.018
S2 × SD _R					-0.008 (0.019)	-0.044, 0.029
S3 × SD _R					-0.040 (0.019)*	-0.077, -0.004
S4 × SD _R					-0.034 (0.016)*	-0.065, -0.004
S5 × SD _R					-0.022 (0.019)	-0.058, 0.014
S6 × SD _R					-0.014 (0.019)	-0.051, 0.022
S7 × SD _R					-	-
μ _u	-0.304 (0.137)*	-0.579, -0.029	-0.332 (0.148)*	-0.629, -0.036		
S1 × μ _u	0.091 (0.098)	-0.102, 0.283	0.088 (0.104)	-0.114, 0.292		
S2 × μ _u	0.204 (0.098)*	0.012, 0.397	0.207 (0.104)*	0.003, 0.410		
S3 × μ _u	0.370 (0.098)***	0.178, 0.563	0.421 (0.104)***	0.217, 0.625		
S4 × μ _u	0.355 (0.082)***	0.194, 0.515	0.431 (0.082)***	0.261, 0.601		
S5 × μ _u	0.234 (0.098)*	0.040, 0.425	0.276 (0.104)**	0.073, 0.480		
S6 × μ _u	0.100 (0.098)	-0.092, 0.292	0.131 (0.104)	-0.072, 0.335		
S7 × μ _u	-	-	-	-		
Bias × μ _u			-0.007 (0.012)	-0.031, 0.018		
S1 × Bias × μ _u			-0.008 (0.009)	-0.018, 0.016		
S2 × Bias × μ _u			0.000 (0.009)	-0.016, 0.017		
S3 × Bias × μ _u			0.123 (0.009)	-0.004, 0.029		
S4 × Bias × μ _u			0.0187 (0.007)**	0.005, 0.033		
S5 × Bias × μ _u			0.011 (0.009)	-0.006, 0.027		
S6 × Bias × μ _u			0.008 (0.009)	-0.009, 0.024		
S7 × Bias × μ _u			-	-		
Bias × SD _R					0.001 (0.002)	-0.003, 0.004
S1 × Bias × SD _R					-0.001 (0.001)	-0.003, 0.002
S2 × Bias × SD _R					-0.002 (0.001)	-0.005, 0.001
S3 × Bias × SD _R					-0.003 (0.001)*	-0.006, -0.001
S4 × Bias × SD _R					-0.002 (0.001)	-0.004, 0.001
S5 × Bias × SD _R					-0.002 (0.001)	-0.005, 0.001
S6 × Bias × SD _R					-0.001 (0.001)	-0.003, 0.002+

S7 × Bias × SD _R			-		-
Random effects					
Variance RI	1.368 (0.327)	-	1.430 (0.350)	-	1.433 (0.3510)
Residual variance	4.690 (0.114)	-	4.67 (0.114)	-	4.707 (0.115)

RI = Random Intercepts , CI = Confidence Interval, SD_R = standard deviation during recall, S7 serves as reference, μ_u = mean uncertainty rating ^ $\leq .1$, * $\leq .05$, ** $\leq .01$, *** $\leq .001$

Day 2 - Startle						
	Model 1		Model 2		Model 3	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	7.481 (1.228)***	5.060, 9.903	7.511 (1.228)***	5.088, 9.933	7.523 (1.230)***	5.097, 9.950
Trial	-0.142 (0.018)***	-0.177, -0.107	-0.142 (0.018)***	-0.177, -0.107	-0.143 (0.018)***	-0.177, -0.108
Stimulus (S1)	-1.034 (1.021)	-3.037, 0.969	-1.018 (1.017)	-3.014, 0.977	-1.034 (1.021)	-3.038, 0.969
Stimulus (S2)	0.084 (1.031)	-1.938, 2.107	0.083 (1.028)	-1.934, 2.100	0.087 (1.031)	-1.936, 2.110
Stimulus (S3)	-0.040 (1.023)	-2.047, 1.967	-0.015 (1.020)	-2.016, 1.986	-0.051 (1.023)	-2.059, 1.957
Stimulus (S4 = CS)	-0.580 (0.884)	-2.315, 1.155	-0.574 (0.881)	-2.303, 1.155	-0.586 (0.885)	-2.321, 1.150
Stimulus (S5)	0.001 (1.019)	-1.998, 2.000	-0.004 (1.016)	-1.996, 1.989	0.009 (1.019)	-2.010, 1.991
Stimulus (S6)	0.043 (1.027)	-1.972, 2.058	0.053 (1.024)	-1.954, 2.061	0.125 (1.028)	-1.893, 2.142
Stimulus (S7)	-	-	-	-	-	-
Bias			0.060 (0.071)	-0.081, 0.200		
S1 × Bias			-0.072 (0.067)*	-0.305, -0.040		
S2 × Bias			-0.008 (0.068)	-0.142, 0.125		
S3 × Bias			0.004 (0.067)	-0.128, 0.136		
S4 × Bias			0.046 (0.058)	-0.068, 0.160		
S5 × Bias			-0.052 (0.67)	-0.183, 0.080		
S6 × Bias			0.008 (0.068)	-0.125, 0.141		
S7 × Bias			-	-		
SD _R					0.062 (0.115)	-0.165, 0.289
S1 × SD _R					-0.062 (0.111)	-0.278, 0.155
S2 × SD _R					-0.033 (0.111)	-0.250, 0.184
S3 × SD _R					-0.043 (0.111)	-0.260, 0.174
S4 × SD _R					0.035 (0.095)	-0.152, 0.222
S5 × SD _R					0.052 (0.111)	-0.163, 0.266
S6 × SD _R					0.165 (0.111)	-0.059, 0.388
S7 × SD _R					-	-
Random effects						
Variance RI	22.619 (5.778)	-	22.837 (5.902)	-	22.782 (5.893)	-
Residual variance	84.901 (3.365)	-	84.277 (3.348)	-	84.934 (3.375)	-

RI = Random Intercepts, CI = Confidence Interval, SD_R = standard deviation during recall, S7 serves as reference, μ_u = mean uncertainty rating $\wedge \leq .1$, * $\leq .05$, ** $\leq .01$, *** $\leq .001$

Day 2 - Startle						
	Model 4		Model 5		Model 6	
Fixed effects	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI	Par. Est. (SE)	95% CI
Intercept	7.490 (1.232)***	5.059, 9.921	7.541 (1.220)***	5.136, 9.948	7.488 (1.229)***	5.059, 9.921
Trial	-0.142 (0.018)***	-0.177, -0.107	-0.142 (0.018)***	-0.177, -0.107	-0.142 (0.018)***	-0.177, -0.107
Stimulus (S1)	-1.046 (1.022)	-3.052, 0.959	-1.045 (1.018)	-3.043, 0.953	-0.975 (1.013)	-2.963, 1.012
Stimulus (S2)	0.063 (1.032)	-1.963, 2.089	0.108 (1.030)	-1.913, 2.128	0.214 (1.026)	-1.798, 2.227
Stimulus (S3)	-0.083 (1.024)	-2.094, 1.928	-0.060 (1.021)	-2.064, 1.943	0.018 (1.016)	-1.975, 2.012
Stimulus (S4 = CS)	-0.590 (0.886)	-2.329, 1.148	-0.589 (0.883)	-2.320, 1.142	-0.682 (0.878)	-2.406, 1.040
Stimulus (S5)	0.001 (1.020)	-2.002, 2.003	-0.003 (1.017)	-1.998, 1.992	-0.004 (1.012)	-1.989, 1.982
Stimulus (S6)	0.073 (1.029)	-1.946, 2.091	0.076 (1.026)	-1.936, 2.089	0.145 (1.012)	-1.856, 2.146
Stimulus (S7)	-	-	-	-	-	-
Bias			0.080 (0.074)	-0.068, 0.277	0.100 (0.075)	-0.049, 0.250
S1 × Bias			-0.130 (0.073)	-0.273, 0.013	-0.217 (0.071)**	-0.357, -0.077
S2 × Bias			0.034 (0.072)	-0.108, 0.176	-0.064 (0.072)	-0.206, 0.078
S3 × Bias			0.008 (0.072)	-0.133, 0.150	-0.038 (0.071)	-0.178, 0.102
S4 × Bias			0.048 (0.062)	-0.074, 0.169	0.065 (0.061)	-0.056, 0.185
S5 × Bias			-0.057 (0.072)	-0.198, 0.084	-0.070 (0.071)	-0.211, 0.069
S6 × Bias			0.014 (0.074)	-0.131, 0.159	-0.035 (0.071)	-0.175, 0.105
S7 × Bias			-	-	-	-
SD _R					-0.008 (0.125)	-0.257, 0.240
S1 × SD _R					0.009 (0.118)	-0.223, 0.240
S2 × SD _R					0.060, (0.118)	-0.172, 0.292
S3 × SD _R					0.040 (0.121)	-0.197, 0.277
S4 × SD _R					-0.004 (0.102)	-0.205, 0.197
S5 × SD _R					0.085 (0.119)	-0.147, 0.318
S6 × SD _R					0.244 (0.121)*	0.006, 0.482
S7 × SD _R					-	-
μ _u	0.082 (0.673)	-1.252, 1.416	0.340 (0.713)	-1.075, 1.755		
S1 × μ _u	0.004 (0.662)	-1.294, 1.302	0.353 (0.699)	-1.019, 1.753		
S2 × μ _u	0.376 (0.663)	-0.925, 1.678	0.676 (0.689)	-0.676, 2.028		
S3 × μ _u	-0.116 (0.641)	-1.373, 1.141	-0.092 (0.680)	-1.427, 1.242		
S4 × μ _u	0.115 (0.571)	-1.005, 1.235	0.147 (0.560)	-1.030, 1.323		
S5 × μ _u	0.670 (0.654)	-0.617, 1.953	0.640 (0.691)	-0.716, 1.996		
S6 × μ _u	0.720 (0.640)	-0.531, 1.972	0.785 (0.677)	-0.542, 2.113		
S7 × μ _u	-	-	-	-		
Bias × μ _u			0.050 (0.056)	-0.062, 0.161		
S1 × Bias × μ _u			0.082 (0.055)	-0.027, 0.190		
S2 × Bias × μ _u			0.099 (0.056)	-0.010, 0.207		
S3 × Bias × μ _u			0.008 (0.053)	-0.098, 0.113		
S4 × Bias × μ _u			0.003 (0.047)	-0.088, 0.095		
S5 × Bias × μ _u			-0.008 (0.054)	-0.115, 0.098		
S6 × Bias × μ _u			0.019 (0.054)	-0.088, 0.125		
S7 × Bias × μ _u			-	-		
Bias × SD _R					-0.014 (0.009)	-0.032, 0.004
S1 × Bias × SD _R					0.015 (0.009)^	-0.002, 0.032
S2 × Bias × SD _R					0.018 (0.009)*	0.001, 0.036
S3 × Bias × SD _R					0.014 (0.009)^	-0.003, 0.031
S4 × Bias × SD _R					-0.004 (0.007)	-0.019, 0.010
S5 × Bias × SD _R					0.008 (0.009)	-0.009, 0.025
S6 × Bias × SD _R					0.017 (0.009)^	-0.001, 0.035
S7 × Bias × SD _R					-	-
Random effects						
Variance RI	22.958 (5.937)	-	84.366 (3.368)	-	83.500 (3.333)	-
Residual variance	85.076 (3.380)	-	21.955 (5.853)	-	23.169 (6.140)	-

RI = Random Intercepts , CI = Confidence Interval, SD_R = standard deviation during recall, S7 serves as reference, μ_u = mean uncertainty rating ^ ≤ .1, * ≤ .05, ** ≤ .01, *** ≤ .001

Day 1

Size estimation

The diameter of the CS was estimated to be 70.50 mm ($SD = 31.52$) on average across trials (Fig. S1A), and estimates increased across trials [trial effect: $F(1,733.02) = 8.32$, $p = .004$, $\theta = .18$, $SE = .06$].

Fig. S1. (A-C) Mean CS size estimate during acquisition, recall and generalization (in mm). Red line is the group average, the grey line indicated the size of the CS. (D-F) Explorative scatterplots between estimated and recalled CS size. (G) Averaged size estimates during the generalization phase across participants (red) and per participants (black). The dotted line indicates the true physical size of the stimuli. (H) Frequency chart of the retrospective number estimate of the number of different CSs. Error bars denote standard errors. S = generalization stimuli, CS = conditioned stimulus.

Day 2

Recall phase

CS count

The bulk of participants reported to be exposed on Day 1 to either one or two stimuli whereas a small subgroup thought many more stimuli were presented on day 1 (see Fig. S1H).

Correlations

In exploratory analyses, we calculated the correlations between the mean CS estimated size (separately for acquisition and generalization) and the mean recalled CS (Fig. S1D-F). CS size estimates on both days strongly correlated ($r = .928, p < .001$) and both did not correlate with recalled CS size ($r_{acq} = .040, p = .804$; $r_{gen} = -.102, p = .535$). Furthermore, there was no correlation between US expectancy ratings on the last trial of acquisition and mean retrospective US expectancy ratings ($r = .155, p = .340$). Finally, we calculated correlations between the amount of variance in estimated CS size across acquisition, recall uncertainty and recall variation. Variation in estimated CS size did not correlate with recall uncertainty ($r = .029, p = .859$) or recall variation ($r = .079, p = .627$), neither was there a relationship between both indices of memory uncertainty ($r = .016, p = .920$). All reported p -values are uncorrected for multiple testing.

Generalization phase

Size estimates

Size estimations increased as stimuli became larger [stimulus effect: $F(6,3231) = 670,981, p < .001$] (Fig. S1G) and did not vary across trials [$F(1,3231) = 0.123, p = .726$].

US-expectancy ratings - Mixed model with 2nd order polynomial (stimulus and stimulus²) using stimulus as continuous predictor

	Model1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)
Intercept	2.693 (.255)***	2.695 (.255)***	2.697 (.255)***	2.689 (.256)***	2.698 (.258)***	2.695 (.259)***
stim	1.274 (.089)***	1.274 (.089)***	1.273 (.088)***	1.273 (.089)***	1.272 (.088)***	1.266 (.089)***
stim ²	-0.154 (.011)***	-0.154 (.011)***	-0.154 (.011)***	-0.154 (.011)***	-0.154 (.011)***	-0.153 (.011)***
Trial	-0.009 (.006)	-0.010 (.007)	-0.010 (.007)	-0.010 (.007)	-0.010 (.007)	-0.010 (.007)
Bias		-.008 (.017)			-.018 (.018)	-.010 (.018)
Bias * stim		.003 (.006)			.010 (.006)	.007 (.007)
Bias * stim ²		.000 (.001)			-.001 (.001)	-.001 (.001)
μ_u			-.432 (.158)***		-.534 (.170)	
μ_u * stim			.254 (.057)***		.320 (.060)***	
μ_u * stim ²			-.0345 (.007)***		-.045 (.007)***	
σ_r				-.027 (.028)		.031 (.030)
σ_r * stim				-.014 (.010)		-.021 (.010)
σ_r * stim ²				.002 (.001)		.003 (.001)*
Bias* μ_u					-.025 (.014)	
Bias* μ_u * stim					.0164 (.005)***	
Bias* μ_u * stim ²					-.002 (.001)***	
Bias* σ_r						.001 (.001)
Bias* σ_r * stim						-.001 (.001)
Bias* σ_r * stim ²						.0002 (.000)*

Par. Est. = parameter estimate, SE = standard error, * = $p \leq .1$, * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$

Startle eye blink responses - Mixed model with 2nd order polynomial (stimulus and stimulus²) using stimulus as continuous predictor

	Model1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)	Par. Est. (SE)
Intercept	6.58 (1.47)***	6.64 (1.47)***	6.58 (1.47)***	6.60 (1.47)***	6.65 (1.46)***	6.77 (1.47)***
stim	.227 (.595)	.223 (.593)	.228 (.595)	.227 (.595)	.229 (.593)	.165 (.593)
stim ²	-.014 (.073)	-.014 (.073)	-.015 (.073)	-.015 (.073)	-.015 (.073)	-.008 (.073)
Trial	-.141 (.018)***	-.141 (.018)***	-.141 (.018)***	-.141 (.018)***	-.141 (.018)***	-.143 (.018)***
Bias		-.174 (.089) [°]			-.077 (.095)	-.214 (.095)*
Bias * stim		.111 (.039)**			.083 (.042) [°]	.140 (.042)***
Bias * stim ²		-.012 (.005)*			-.009 (.005) [°]	-.014 (.005)**
μ_u			-.102 (.864)		.753 (.913)	
μ_u * stim			.179 (.389)		-.045 (.410)	
μ_u * stim ²			-.015 (.047)		.005 (.050)	
σ_r				-.083 (.145)		-.027 (.160)
σ_r * stim				.068 (.065)		.025 (.071)
σ_r * stim ²				-.006 (.008)		-.001 (.008)
Bias* μ_u					.200 (.073)**	
Bias* μ_u * stim					-.056 (.033) [°]	
Bias* μ_u * stim ²					.005 (.003)	
Bias* σ_r						.011 (.011)
Bias* σ_r * stim						-.008 (.005)
Bias* σ_r * stim ²						.001 (.001)

Par. Est. = parameter estimate, SE = standard error, [°] = $p \leq .1$, * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$

Effects of STAI-T and PSQ

No effects of STAI-T or GSQ scores assessed on Day 2 were found on the shape of the US expectancy generalization gradient [STAI-T effect: $F(1,44) < 1$, $p = .902$; STAI-T $\times$ stimulus effect: $F(6,3357) < 1$, $p = .500$; GSQ effect: $F(1,44) < 1$, $p = .537$; GSQ $\times$ stimulus effect: $F(6,3357) < 1$, $p = .457$].

No effects of STAI-T or GSQ scores were found on the shape of the gradient for startle eye blink responses apart from a main effect of STAI-T with less potentiation for higher STAI-T scores ($\beta = -.204$, $SE = .11$) [STAI-T effect: $F(1,38) = 5.419$, $p = .025$; STAI-T $\times$ stimulus effect: $F(6,1269) = 1.233$, $p = .286$; GSQ effect: $F(1,38) = 1.487$, $p = .230$; GSQ $\times$ stimulus effect: $F(6,1270) < 1$, $p = .627$].