

**Main Manuscript for**

Threat extinction without awareness: evidence from

masking and continuous flash suppression

Shani Bachar-Avnieli1\*, Liad Mudrik2, 3,. Daniela Schiller4, Gideon E. Anholt1

1Department of Psychology, Ben-Gurion University of the Negev, Beer-Sheva, Israel.

2Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.

3School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel. Masking, Unconscious processes, Anxiety.

**Author Contributions**

**This PDF file includes:**

Main Text

Figures 1 to 5

**Abstract**

Anxiety disorders are the most common mental disorders. The typical treatment for these disorders is exposure to the anxiety inducer. However, many patients are reluctant to confront feared objects or situations and are consequently left untreated. To date, studies showed that threat could be acquired without awareness, though there is no conclusive laboratory evidence that threat can be extinguished unconsciously. The present research aims to evaluate the feasibility and robustness of extinction evoked by unconsciously perceived stimuli, using two common methodologies: Continuous Flash Suppression (CSF) and Visual Masking (VM). We describe two experiments utilizing these two methods (N=41 and N=72, respectively). In each experiment, healthy participants were first threat-conditioned, and later randomly divided into three groups: The first underwent conscious exposure to the threat-conditioned stimuli in order to evoke extinction; the second was unconsciously exposed to these stimuli; and the third served as a control group that did not undergo extinction. In both experiments, we observed effective extinction and successful extinction retention (indexed by a skin conductance response) under both conscious and unconscious exposure conditions. These results demonstrate the effectiveness of unconscious extinction, even when strict measures of awareness are taken. The convergence of both the VM and the CFS techniques bears theoretical implications for the understanding of exposure therapy and may pave a path for the potential clinical utility of unconscious extinction.

**Significance Statement**

Exposure is an effective psychological intervention for the treatment of anxiety. However, this method is associated with considerable treatment refusal and high drop-out rates. Thus, there is a need for alternative treatment methods that will help patients minimize the encounter with aversive stimuli. A possible advancement to this problem may be found in the field of consciousness studies, where different techniques were developed to present stimuli so that they will be processed without conscious awareness. The current research goal is to test for extinction while assessing conscious experience of the suppressed stimuli by using two methodologies, CFS and VM.

The most common intervention for anxiety disorders is treatment with exposure to a feared situation or object (1). The goal of exposure is to facilitate extinction – that is, to reduce the conditioned threat response to the triggering stimuli (2). Although effective psychological and pharmacological treatments exist for anxiety disorders (3), most people with such disorders never seek treatment (4). A possible explanation for these low rates of treatment-seeking is that patients consider confronting feared objects or situations as overly demanding. Accordingly, novel interventions that will help patients minimize their encounter with aversive stimuli are needed (5). A possible advancement towards this goal may be found in the field of consciousness studies, where different techniques have been developed to present stimuli so that they will be processed without conscious awareness. The current research aimed at exploring the feasibility of unconscious extinction.

In recent years, there has been increasing evidence suggesting that threat responses can be acquired or evoked by stimuli that are rendered invisible (6-9). For example, one study reported that fearful facial expressions emerge from suppression into awareness more quickly relative to images of neutral or happy expressions, suggesting that emotional expressions might be unconsciously processed (10). These results are supported by findings demonstrating increased amygdala activation in response to fearful, compared with happy, masked faces (but see (11)). These findings are further in line with LeDoux’s (12) suggestion of a direct path between the thalamus and the amygdala, such that information may evoke defensive responses, even without activation of the visual cortex.

If threat can be acquired and experimentally evoked outside of awareness, it is plausible that conditioned threat could also undergo extinction under similar conditions. However, previous literature does not provide clear and compelling evidence for extinction evoked by invisible stimuli. Several behavioral studies tested whether subliminal exposure to images of spiders affects one’s willingness to approach a spider among people who are afraid of spiders (13, 14). In these studies, participants completed a behavioral avoidance test (BAT) one week after a masked exposure to a spider to examine long-term effects of this form of exposure. They were then presented with images of spiders, either masked or unmasked. Participants in the masked condition were more willing to approach a spider than those who were consciously exposed to pictures of spiders. These findings were replicated with two-week and one-year follow-up measurement (14, 15). However, these behavioral effects have yet to be corroborated by physiological responses. A recent study measured participants’ skin conductance in response to exposure and concluded that masked exposure is not associated with increased physiological responses in the extinction process. Most importantly, although participants in the masked condition did succeed more in the BAT, no evidence for reduced physiological responses was obtained (16). Yet the most critical point regarding the above findings refers to the inability to determine if the participants were indeed unaware of the extinction stimuli. In these studies (13-16), no online measures of awareness (neither subjective or objective; see (17)) were taken. Instead, the premisethat participants were not aware of the stimuli was based on a preliminary masking experiment, in which participants were unable to identify the masked images. Of particular interest is recent work showing the potential benefit which relies on unconscious exposure using CFS (18). In that study, fear was reduced through a measure of threat-potentiated startle responses but not in SCR. Furthermore, that study lacked a control group.

Given these gaps, our goal was to investigate the effectiveness of unconscious exposure to aversive stimuli, when awareness is properly controlled for. Various techniques have been developed to suppress stimuli from awareness and assess unconscious processing. These techniques were used to measure the impact of specific stimuli on participants’ thoughts, feelings, actions and learning processes (19, 20). We focused on two prominent techniques: VM (21, 22) and CFS (23). In VM, a stimulus (“target”) is presented for a short duration of several dozens of milliseconds or less, and is immediately preceded/followed by masks, which render it invisible (21, 22, 24). CFS relies on dichoptic vision: A target stimulus is presented to one eye, while the other eye is consistently exposed to a changing pattern of different shapes. This technique prevents participants from seeing the constant target image for a relatively long period of time (up to several seconds) (25). Importantly, the two techniques may involve different underlying mechanisms and may evoke different types of unconscious processing (22, 26, 27). Therefore, employing both unconscious methodologies is advantageous when studying unconscious processes, and provides means to assess the robustness and generalizability of the effect. To assess the threat response, we relied on a commonly used autonomic measure of threat conditioning – the Skin Conductance Response (SCR) (9). Studies have demonstrated that this measure contributes to the understanding of anxiety disorders when used in experiments of fear conditioning (28).

The present research aims at evaluating the feasibility and robustness of extinction evoked by unconsciously perceived stimuli by (a) testing for extinction while assessing conscious experience of the suppressed stimuli; and (b) testing for the robustness and generalizability of non-conscious extinction across both CFS (Experiment 1) and VM (Experiment 2).

Both experiments included measurements of changes in skin conductance, while viewing a pre-installed computer presentation on a monitor. The studies started with an acquisition phase, where the participants were presented with Conditioned Stimuli (CS+ and CS-): a scared face of a man or a woman. While the conditioned stimulus was presented, participants received mild electric shocks at a level of which they determine to be “aversive” and “uncomfortable, but not painful”. Subsequently, participants underwent an extinction phase, where they were presented with the same stimuli again, without receiving the electric shock. Participants were divided into 3 groups: (1) “Unaware group”, (2) “Aware group”, (3) “No Extinction group”. Finally, all participants underwent a testing phase to assess the effects of conscious and unconscious extinction relative to the No Extinction group.

**EXPERIMENT 1**

**Is unconscious threat extinction using CSF possible?**

**Results and Discussion**

*Assessment of Awareness*

In this experiment, awareness was assessed according to both subjective and objective criteria. The subjective ratings showed that 80.2% of the trials were rated as 1 (“I did not see anything”), and 14.6 as 2 (“I had a vague perception of something”). Only 4.89% were rated as either 3 (“I saw a clear part of the image”) or 4 (“I saw the picture clearly”). Objective performance showed that in visibility 1 trials, participants’ accuracy in the gender judgment task was 49% (SD=0.11), and did not differ from chance [(t(10) = 0.219, p = 0.830]. This null result was supported by a Bayesian paired-sample t-test which revealed that, given our data, the null hypothesis was 3.14 times more likely than the alternate hypothesis. This indicates that the masking procedures were effective and that participants were unaware of the gender faces. To provide a more convincing test for the unaware manipulation, we used an Even-Odd analysis as recommended by Shanks (29). This entailed the calculation of two independent objective performance measures, one calculated from even-numbered trials, and the other from odd-numbered trials. We found that the two measures were strongly correlated (r (13) = 0.91, P < 0.001), suggesting they are reliable in assessing objective awareness. Taken together, awareness assessments confirmed that participants were not consciously aware of the masked stimuli.

**conditioned threat acquisition**

To evaluate the acquisition process across the assigned groups, we conducted an ANOVA with factors of group (Unaware, Aware, No-extinction) × stimulus (CS+, CS−). There was no evidence for group differences as indicated by the non-significant main effect of group (F(1,38) = 2.64, p=0.084, partial η2 = 0.12) or group × stimulus interaction (F(2,38) = 1.22, p = 0.304, partial η2 = 0.06 ). SCR responses were higher to the CS+ compared to the CS− [F (1,38) =311.8, p < 0.001, partial η2 = 0.89]. As expected, these results confirm sufficient and similar acquisition across all groups (F < 1, ns).

**Threat extinction**

To evaluate extinction learning, as well as potential group differences during fear and extinction learning, we focused on CS difference scores averaged over the extinction phase and checked if they differed between groups and between the early and late phases of extinction, defined as 12 first trials and 12 last trials. Accordingly, CS difference scores were entered into a 3 × 2 repeated measures ANOVA with factors group (Unaware; Aware; no-extinction) and time (Early; Late). Main effects of time [F (1, 35) = 9.479, p = 0.004, partial *η*2 = 0.21]. However, no main effects of group [F (2, 35) = 1.60, p = 0.216, partial *η*2 = 0.08], or interaction of group and time, were found [F (2, 35) = 0.58, p =0.565, partial *η*2 = 0.03].

Pairwise comparison analysis for each group separately showed that for the unaware group, responses marginally significant decrease between Early and Late Extinction Phases (p = 0.07), suggesting that extinction can occur also when the stimuli are rendered invisible. As expected, extinction was found also in the aware group, where responses declined between Early and Late Extinction (p =0.01). As opposed to both the unaware and the aware groups, responses in the no-extinction group did not differ between early and late Extinction (p = 0.328; See Figure 2), as expected.

**Extinction retention**

To evaluate between-group differences in the extinction process, a Recovery Index (RI) was calculated by deducting the CS-elicited threat response at the last four trials of extinction from the trials of the Test Phase. Univariate analysis confirmed a group difference on the RI, yielding an effect size (partial η2) of 0.34, see Figure 3. Where the no-extinction group (M = 0.116 ; p=0.001) showed clear indications of recovery, and no recovery was observed in the aware groups (M = -0.077; p=0.02). In addition, no recovery was observed in the unaware group as well, (M = -0.02), however this difference was not significant (p=0.501).

Altogether, in experiment 1 decreases in SCR were demonstrated among participants in the exposure groups, but not those in the control group. The results strengthen the findings that threat extinction can be learned without conscious awareness. As far as we know, this study is the first to examine unconscious extinction in controlled laboratory settings, by using CFS.

**EXPERIMENT 2**

**Is unconscious threat extinction using VM possible?**

**Results and Discussion**

*Assessment of Awareness*

Akin to the CFS experiment, here too, the vast majority of trials (83.9%) were rated as 1 (“I did not see anything”), and 13.1% as 2 (“I had a vague perception of something”). Only 2.89% were rated as either 3 (“I saw a clear part of the image”) or 4 (“I saw the picture clearly”). Objective performance showed that in visibility 1 trials, participants’ accuracy in the gender judgment task was again at chance (M=48%, SD=??) [ t (23) = -1.04, p =0.152]. Bayesian paired-sample t-test suggested that the null hypothesis was 2.86 times more likely than the alternate hypothesis. Again, Even-Odd analysis suggested a high reliability of the objective measures (r (23) = 0.85, P < 0.001). Thus, much like in Exp. 1, the awareness assessments in Exp. 2 suggested that masking was effective in suppressing the stimuli from awareness.

*Conditioned threat acquisition*

An ANOVA on the CS difference scores with factors of group (Unaware, Aware, No-extinction) × stimulus (CS+, CS−) yielded no evidence for group differences (F(1,69) = 0.34, P= 0.71, partial η2 = 0.01) or group × stimulus interaction (F(2,69) = 0.85, P = 0.43, partial η2 = 0.02), yet SCR responses were higher to the CS+ compared to the CS− [F (1,69) =139.92, p < 0.001, partial η2 = 0.67]. As expected, these results confirm sufficient and similar acquisition across all groups (F < 1, ns).

*Threat extinction*

A 3 × 2 repeated measures ANOVA with factors group (Unaware; Aware; no-extinction) and Phase (Early; Late) yielded no group differences [F (1,69) =0.06, p =9.35, partial η2 = 0.02]. However, main effects of time [F (2, 69) = 89.53, p < 0.001, partial *η*2 = 0.56], as well as the interaction of group and time, were found [F (2, 69) = 25.01, p < 0.001, partial *η*2 = 0.42].

 Pairwise comparison analysis for each group separately clearly replicated the results of Experiment 1: For the unaware group, responses declined between Early and Late Extinction Phases (p < 0.0001). Similarly, in the aware group, responses declined between Early and Late Extinction (p < 0.0001). As expected, in contrast to the unaware and the aware groups, the responses in the no-extinction group did not differ between early and late Extinction (p = 0.515; See Figure 4).

*Extinction Retention*

Univariate analysis again replicated the results of Exp. 1 also with respect to the RI, where the no-extinction group (M = .0608) showed clear indications of recovery, whereas no recovery was observed in the unaware (M = -.0546) and the aware groups (M = -.0467), yielding an effect size (partial η2) of 0.42. The RI differences scores for the three groups (Figure 5) showed clear indications of recovery for the control group, whereas no recovery was observed in the aware and unaware groups.

Thus, Exp. 2 provided a complete replication of the results of Exp. 1, despite using a different technique for rendering stimuli invisible: decreases in SCR were demonstrated among participants in the aware and unaware exposure groups, but not those in the no-extinction group. These results strengthen the findings of Siegal et al (16), where a decrease in SCR was found following exposure to masked stimuli, yet without a trial-by-trial rigorous measurement of awareness.

*Assessment of techniques: CFS Vs. VM*

We used Hedges g’ (30) between group effect size to assess the differences between the unaware group and the no-extinction group in both CFS and VM techniques. Effect sized were obtained by SCR of the Recovery index in the unaware group that went through the unconscious extinction, was compared with the no-extinction group.

The effect size was calculated using the following formula: $g$= $\overbar{X}$t+$\overbar{X}$c /Sp, where $\overbar{X}$t is the mean of the unaware group, $\overbar{X}$c is the mean of the no-extinction group, and Sp is the pooled standard deviation. The CFS technique in the first experiment showed an effect size of hedge’s g: 1.16 (95% CI: 0.308-2.013), whereas the VM technique in the second experiment showed an effect size of hedge’s g: 2.09 (95% CI: 1.394-2.803). This means that the effect size of unconscious exposure in the VM technique was larger than the effect size of the CFS group technique.

**General Discussion**

Using two different methods to suppress stimuli from awareness, CFS and VM, we demonstrated that unconscious extinction can occur. Furthermore, the effect of unconscious exposure in the VM technique, was larger than the CFS technique. Importantly, all experimental groups showed comparable threat-response in the acquisition process, with higher SCR to the CS+ relative to the CS- stimulus. However, in the late extinction in both CFS and VM, only the aware and the unaware groups showed decreased SCR. This result was strengthened by the RI measure, demonstrating that recovery occurred only among participants that did not undergo extinction; there was no recovery evident among participants in the unaware group or the aware groups.

Findings from both experiments suggest that conscious and nonconscious exposure is effective in reduction through non-reinforced presentations. These findings could pave the way for establishing a new therapeutic protocol, which relies on unconscious exposure. Such a protocol has never been implemented; however, some studies suggest that extinction through unconscious exposure may be effective. There is a growing body of research on a novel brain imaging approach called decoded fMRI neurofeedback (31-33) This technique depends on rewarding unconscious neural representations of feared stimuli in order to counter-condition the feared representation. This approach has demonstrated promise in decreasing fear responses to laboratory conditioned fears (34). Both approaches are similar in showing that fear extinction can be learned without consciousness. Note however, that the approach untilized in the current research is much more easy to implement and does not require complex infrastructure.

Our finding that extinction might be independent from awareness has interesting theoretical implications. One of the predominant theories of exposure therapy is the inhibitory learning model (35' 36). This model suggests that the relationship between the CS and the aversive stimulus is not eradicated. Rather, a new inhibitory connection is created, whereby the conditioned stimulus no longer predicts the aversive stimulus (thereby inhibiting the fear response). The inhibitory connection then “competes” with the previous fear learning. One of the core processes suggested to underly inhibitory learning is expectancy violation. It is based on the premise that a gap between expectations and actual outcomes is critical for learning new inhibitory expectations, competing with previous expectations. However, given that extinction learning is based on the formation of non-coincidental relationships between conditioned and unconditioned stimuli, awareness of the stimuli as well as the non-occurrence of the unconditioned stimulus is deemed essential (37). The current findings indicate that extinction can also occur outside of conscious awareness, which goes against this assumption. Future research may indicate whether unconscious exposure involves expectancy violation or other mechanisms (e.g., habituation).

In the current study, we chose to use two techniques to render stimuli invisible: CSF and VM. In the CFS technique, each eye is presented with a different stimulus. One eye is presented with a series of flashing high contrast images, while the other eye is presented with a stationary, often low-contrast target. In the VM technique, the target stimulus is displayed for a short time period of several dozens of milliseconds and is immediately preceded and followed by a masked stimulus. The CFS technique is based on the fact that the visual system is not able to handle incompatible input to both eyes. As a result, only one stimulus is able to reach awareness, while the other stimulus remains invisible (until it breaks suppression; (10)). In contrast, in the VM technique, the mask stimuli tampers with the feedback sweep that typically follows feedforward processing of the target stimulus (22). As such, the differences between the two techniques also influence the amount of information that gets through and is being processed (38). Despite the differences between the techniques (27, 39), in our study both techniques evoked successful extinction. Nevertheless, the VM technique outperformed the CFS. Studies have shown that unconscious processing under CFS is more limited than VM in presenting unconscious semantic effects (40, 41). Other studies indicate that backward masking represents a more sensitive technique for measuring unconscious high-level processing than interocular suppression (19).

The present study has a number of limitations and, as such, our findings should be interpreted with caution. The present study used immediate exposure, so that consolidation of the previously acquired fear memory would not be possible. Future studies should utilize a 3-day study design, in which the acquisition phase would be separated from the extinction phase by at least 24 h. In conclusion, the present study demonstrated that unconscious exposure using both CFS and VM, and that the effect size of extinction using VM is larger.

Results may facilitate the development of novel treatments integrating unconscious exposure with various psychopathologies, populations, therapeutic doses and combinations between unconscious and traditional in vivo treatment regiments.

**Method**

**Experiment 1**

**Is unconscious threat extinction using CSF possible?**

*Participants*

Forty-one healthy participants with normal or corrected normal visual acuity from the department of Psychology participated in the current experiment for course credit (32 women, 38 right-handed, mean age = 25, SD = 1.33). Seven participants who did not have measurable responses to the shocks were not included in the data analysis (SCR Score > 0.2). The experiment was approved by the ethics Committee of the Ben-Gurion University. Participants with self-reported psychiatric or neurologic history were excluded. All participants signed an informed consent form before partaking in the experiment.

*Stimuli and apparatus*

Two grayscale photographs of a woman and a man with identical contrast and luminance degree, each expressing fear, served as the target stimuli in all groups. They were selected from the NIMSTIM Database (42). The choice to present fearful facial expressions was based on previous studies (7, 9, 11) showing that an angry or scared face can be identified faster than a smiling face, and that a fearful face might indicate a potential danger and is accordingly more easily conditioned.

The stimuli were presented on a black background. The pictures were blurred at the tips by using Photoshop software and were surrounded by black-and-white rectangle frames, as depicted in Figure 1, to facilitate fusion. The experiment was computerized, performed on a 19-inch color Samsung screen with 60 HZ refresh rate and 1024 × 768 resolution, using E-Prime software version 2.0. Participants performed the experiment in a slightly darkened room. Their heads were maintained 61 cm from the screen using a chin-rest. The stereoscope was produced by Stereo Aids (Western Australia), and allowed the presentation of a separate image to each eye. Electric shocks were administered to the participant via STIMSOLA system of Biopac Company. The system includes a STMISOLA slider and a USB component which enabled the communication between the shocker appliance and the E-Prime software. The power of the electric shock was in the range of 0–50 v and the shock’s duration was 200 ms. A snap electrode with isotonic gel was attached to the participants’ arms.

Fear arousal was measured through SCR, using the 150 MP system of the Biopac GSR100C Company. Samples were collected with the Acknowledge system of Biopac Company.

*Procedure*

All participants underwent three phases during the experiment: acquisition, extinction, and testing within one day (see Figure 1). The experiment started with an acquisition phase, which was done using a threat conditioning paradigm with partial reinforcement. A male/female face served as the conditioned stimulus (CS+; counterbalanced between subjects), and an electric shock was used as the unconditioned stimulus (US). The acquisition phase included 24 trials: four practice trials, 12 CS+ trials (e.g., female face presented alone), 12 CS- trials (male face presented alone), six CS-US trials (female face presented with a shock. The order of stimuli appearance was pseudo-randomized: the first trial was always reinforced and no more than two of the same trial types ever occurred consecutively. Subsequently, participants underwent an extinction phase, in which they were presented with the same stimuli again, but without receiving electric shock (12 CS+ trials, 12 CS- trials). In the extinction phase, participants were divided into three groups: (1) The *unaware group* )n=13) was presented with the CS+/CS- stimuli under CFS, so the stimuli were presented to one of their eyes, while a flickering stimulus of colored squares (Mondrian) was presented to their other eye, (2) the *aware group* )n=14) was presented with the CS+/CS- stimuli without CFS, (3) the no-extinction group )n=14) was presented with scrambled versions of the CS+/CS- stimuli under CFS. In all groups, the stimuli were presented for 4 s . Participants in the unaware group and the no-extinction group, in which CFS was used, were asked two questions at the end of each trial in order to confirm their level of awareness to the stimulus. These questions were used as objective and subjective measures to assess the level of the participants’ awareness. The objective question was: “Was the person in the picture a man or a woman?” The subjective question was: “How visible was the picture?”, and subjects were asked to respond using the Perceptual Awareness Scale (PAS:(43)), where 1 denotes “I did not see anything”; 2 represents “I had a vague perception of something”; 3 signifies “I saw a clear part of the image”; and 4 indexes “I saw the picture clearly”. Finally, in testing phase to assess the effects of conscious and unconscious extinction were assessed relative to the no-extinction group. In the testing phase, all three groups were presented with the CS+/CS- stimuli, in order to measure the recovery process in all three groups.

*Psychophysiological stimulation and assessment*

Participants viewed a pre-installed computer presentation on a monitor while changes in their SCR were measured. For SCR recordings, electrodes were attached to the forefinger and the forearm on each participant’s left hand While SCR was measured, participants received mild electric shocks at a level that they determined to be “aversive” (undesired and unpleasant) and “uncomfortable, but not painful” (25). The electric shock appeared randomly about 0.5–4.5 s from the moment that the stimulus was presented. Between one stimulus and another, there was a time gap of 8–12 s. The electric shock was delivered when participants were presented with a CS+ stimulus, but never when presented with a CS- stimulus.

*SCR Analysis*

SCR waveforms were analyzed offline, using the Acknowledge 3.9 software (Biopac Systems Inc.). SCR amplitudes to the conditioned and unconditioned stimuli were the dependent measures of conditioned and unconditioned responses, respectively. The level of SCR was determined by taking the base-to-peak difference for the first waveform (in microsegments, ms) in the 0.5–4.5 s window after stimulus onset. The minimal response criterion was 0.02 ms. The raw SCR scores were square-root-transformed to normalize distributions. These normalized scores were scaled according to each participant’s unconditioned response by dividing each response by the mean square-root-transformed unconditioned stimulus response.

**Experiment 1**

**Is unconscious threat extinction using VM possible?**

The methods of experiment 2 were identical to those described in experiment 1, with the following exceptions:

*Participants*

The GPower software version 3.0.5 (29) was used to determine the required sample size to obtain an effect size (ES) estimate of 0.25, which was chosen based on the results of the first experiment. The projected sample size needed for an ES of 0.25 with an alpha of 0.05, power (1 − β) of 0.95, three groups, and three repetitions was 60. As such, we decided to recruit 72 participants for the current study (53 women, 65 right-handed, mean age = 25, SD = 1.93). Eleven additional participants also completed the experiment but were excluded for the following reasons: three participants were excluded due to technical problems with data recording, eight participants were classified as non-responders because they lacked a measurable SCR (defined as a response higher than 0.02 ms) on >75% of the trials and, thus, were excluded from analyses.

*Procedure*

Similarity to experiment 1, participants first underwent a fear acquisition. In the second phase, participants were randomly assigned to one of three extinction groups: (1) “unaware group”- a subliminal stimulus (face of a man or a woman, as detailed in experiment 1) was presented by VM, (2) “aware group”- a face of a man or a woman was presented for a duration of 4 s, and (3) “no-extinction group”- a scrambled face stimuli was presented by VM. The no-extinction group did not undergo extinction. The target stimulus was presented for 33 ms and the masked stimuli were presented for an additional 6 s. Finally, all participants underwent a testing phase to assess the effects of conscious and unconscious extinction relative to the no-extinction group.

**Acknowledgments**

**References**

1. B. J. Deacon, J. S. Abramowitz, Cognitive and behavioral treatments for anxiety disorders: A review of meta‐analytic findings. *Journal of clinical psychology* **60**, 429-441 (2004).

2. J. S. Abramowitz, The practice of exposure therapy: relevance of cognitive-behavioral theory and extinction theory. *Behavior therapy* **44**, 548-558 (2013).

3. E. M. Investigators *et al.*, Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta psychiatrica scandinavica* **109**, 21-27 (2004).

4. H.-U. Wittchen *et al.*, The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European neuropsychopharmacology* **21**, 655-679 (2011).

5. J. Pearson, Associative learning: Pavlovian conditioning without awareness. *Current Biology* **22**, R495-R496 (2012).

6. C. M. Raio, D. Carmel, M. Carrasco, E. A. Phelps, Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology* **22**, R477-R479 (2012).

7. A. Öhman, S. Mineka, Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological review* **108**, 483 (2001).

8. A. Öhman, Fear and anxiety as emotional phenomena: clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. (1993).

9. F. Esteves, C. Parra, U. Dimberg, A. Öhman, Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear‐relevant facial stimuli. *Psychophysiology* **31**, 375-385 (1994).

10. T. Stein, M. V. Peelen, Dissociating conscious and unconscious influences on visual detection effects. *Nature Human Behaviour*, 1-13 (2021).

11. P. J. Whalen *et al.*, Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience* **18**, 411-418 (1998).

12. J. LeDoux, The emotional brain: The mysterious underpinnings of emotional life. *World and I* **12**, 281-285 (1997).

13. J. Weinberger, P. Siegel, C. Siefert, J. Drwal, What you cannot see can help you: The effect of exposure to unreportable stimuli on approach behavior. *Consciousness and cognition* **20**, 173-180 (2011).

14. P. Siegel, J. Weinberger, Very brief exposure: The effects of unreportable stimuli on fearful behavior. *Consciousness and Cognition* **18**, 939-951 (2009).

15. P. Siegel, R. Warren, Less is still more: Maintenance of the very brief exposure effect 1 year later. *Emotion* **13**, 338 (2013).

16. P. Siegel *et al.*, Less is more: Neural activity during very brief and clearly visible exposure to phobic stimuli. *Human brain mapping* **38**, 2466-2481 (2017).

17. E. M. Reingold, P. M. Merikle, Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics* **44**, 563-575 (1988).

18. J. P. Oyarzún, E. Càmara, S. Kouider, L. Fuentemilla, R. de Diego‐Balaguer, Implicit but not explicit extinction to threat‐conditioned stimulus prevents spontaneous recovery of threat‐potentiated startle responses in humans. *Brain and behavior* **9**, e01157 (2019).

19. P. Sterzer, T. Stein, K. Ludwig, M. Rothkirch, G. Hesselmann, Neural processing of visual information under interocular suppression: a critical review. *Frontiers in psychology* **5**, 453 (2014).

20. S. Kouider, S. Dehaene, Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 857-875 (2007).

21. D. Kahneman, Method, findings, and theory in studies of visual masking. *Psychological Bulletin* **70**, 404 (1968).

22. B. G. Breitmeyer, H. Ogmen, Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & psychophysics* **62**, 1572-1595 (2000).

23. N. Tsuchiya, C. Koch, Continuous flash suppression reduces negative afterimages. *Nature neuroscience* **8**, 1096-1101 (2005).

24. B. G. Breitmeyer, T. Ro, H. Ogmen, A comparison of masking by visual and transcranial magnetic stimulation: implications for the study of conscious and unconscious visual processing. *Consciousness and Cognition* **13**, 829-843 (2004).

25. A. Öhman, G. Erixon, I. Löfberg, Phobias and preparedness: Phobic versus neutral pictures as conditioned stimuli for human autonomic responses. *Journal of Abnormal Psychology* **84**, 41 (1975).

26. C.-Y. Kim, R. Blake, Psychophysical magic: rendering the visible ‘invisible’. *Trends in cognitive sciences* **9**, 381-388 (2005).

27. S. V. Fogelson, P. J. Kohler, K. J. Miller, R. Granger, P. U. Tse, Unconscious neural processing differs with method used to render stimuli invisible. *Frontiers in psychology* **5**, 601 (2014).

28. S. C. Bunce, E. Bernat, P. S. Wong, H. Shevrin, Further evidence for unconscious learning: Preliminary support for the conditioning of facial EMG to subliminal stimuli. *Journal of psychiatric research* **33**, 341-347 (1999).

29. D. R. Shanks, Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review* **24**, 752-775 (2017).

30. R. Rosenthal, H. Cooper, L. Hedges, Parametric measures of effect size. *The handbook of research synthesis* **621**, 231-244 (1994).

31. V. Taschereau-Dumouchel *et al.*, Towards an unconscious neurotherapy for common fears. *bioRxiv*, 170183 (2017).

32. R. Sitaram *et al.*, Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience* **18**, 86-100 (2017).

33. K. Shibata, T. Watanabe, Y. Sasaki, M. Kawato, Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *science* **334**, 1413-1415 (2011).

34. A. Koizumi *et al.*, Fear reduction without fear through reinforcement of neural activity that bypasses conscious exposure. *Nature human behaviour* **1**, 1-7 (2016).

35. M. E. Bouton, Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological bulletin* **114**, 80 (1993).

36. S. M. Blakey, J. S. Abramowitz, The effects of safety behaviors during exposure therapy for anxiety: Critical analysis from an inhibitory learning perspective. *Clinical Psychology Review* **49**, 1-15 (2016).

37. M. G. Craske, M. Treanor, C. C. Conway, T. Zbozinek, B. Vervliet, Maximizing exposure therapy: An inhibitory learning approach. *Behaviour research and therapy* **58**, 10-23 (2014).

38. B. G. Breitmeyer, Psychophysical “blinding” methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition* **35**, 234-250 (2015).

39. E. J. Cox, I. Sperandio, R. Laycock, P. A. Chouinard, Conscious awareness is required for the perceptual discrimination of threatening animal stimuli: a visual masking and continuous flash suppression study. *Consciousness and cognition* **65**, 280-292 (2018).

40. Y.-H. Yang, Y.-H. Tien, P.-L. Yang, S.-L. Yeh, Role of consciousness in temporal integration of semantic information. *Cognitive, Affective, & Behavioral Neuroscience* **17**, 954-972 (2017).

41. M.-S. Kang, R. Blake, G. F. Woodman, Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience* **31**, 13535-13545 (2011).

42. N. Tottenham *et al.*, The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research* **168**, 242-249 (2009).

43. T. Z. Ramsøy, M. Overgaard, Introspection and subliminal perception. *Phenomenology and the cognitive sciences* **3**, 1-23 (2004).

**Figures and Tables**

****

|  |  |
| --- | --- |
|  |  |
| **Figure 1. Experimental procedure and stimuli**. All participants underwent three phases in the experiment: acquisition, extinction, and test phases. Acquisition phase (left): participants were presented with a male/female face, while receiving an electric shock for some of the trials. Extinction phase (right): participants were again presented with the same stimuli that were either intact or scrambled (in the no acquisition group), without receiving the electric shock. In the aware group, the stimuli were visible, while in the unaware group, depicted here, they were rendered invisible using CFS. Participants were asked two questions at the end of each trial. These questions were used as objective and subjective measures to assess the level of the participants’ awareness. Finally, all participants underwent a testing phase to assess the effects of conscious and unconscious extinction forms relative to the no-extinction group.

|  |  |  |
| --- | --- | --- |
|

|  |
| --- |
|  |
| **Figure 2. Experiment 1 (CFS): Evidence for extinction across the groups.** **Normalized SCRs for early and late extinction.** Extinction was analyzed with a 3\*2 repeated measures of ANOVA with factors of Time (Early Extinction, Late Extinction) and Group (Aware, Unaware, No extinction) which showed a significant main effect of time [F (1, 35) = 9.479, p = 0.004, partial *η*2 = 0.21], but no main effect of Group [F (2, 35) = 1.60, p = 0.216, partial *η*2 = 0.08], or interaction of Group and Time, were found [F (2, 35) = 0.58, p =0.565, partial *η*2 = 0.03]. Pairwise comparison analysis showed no SCR difference between Early and Late Extinction in the No-Extinction (p = 0.328), but responses declined between Early and Late Extinction Phases in the aware group (p =0.01) and marginally significant decrease in unaware groups (p = 0.07).  |

 |

|  |  |
| --- | --- |
| Chart, box and whisker chart  Description automatically generated |  |
|  |  |
| **Figure 3. Experiment 1- CFS-** **indications of recovery.** Recovery index (RI) was defined by deducting the CS elicited threat response at the last four trial of extinction from the trials of the Test Phase. Univariate analysis confirmed a significant group difference on the RI, whereas the no-extinction group (M = 0.116) showed robust recovery, no recovery was observed in the unaware (M = -0.02) and the aware groups (M = -0.077), yielding an effect size (partial η2) of 0.34. |

|  |  |
| --- | --- |
|  |  |
| **Figure 4. Experiment 2 (Visual Masking): Evidence for extinction across the groups.** **Normalized SCRs for early and late extinction.** Extinction was analyzed with a 3\*2 repeated measures of ANOVA with factors of Time (Early Extinction, Late Extinction) and Group (Aware, Unaware, No extinction) which showed a significant main effect of [F (2, 69) = 89.53, p < 0.001, partial *η*2 = 0.56], but no main effect of Group [F (1,69) =0.06, p =9.35, partial η2 = 0.02], or interaction of Group and Time, were found [F (2, 69) = 25.01, p < 0.001, partial *η*2 = 0.42]. Pairwise comparison analysis showed no SCR difference between Early and Late Extinction in the No-Extinction group (p = 0.515), but responses declined between Early and Late Extinction Phases in the aware and unaware group (p < 0.0001).

|  |  |
| --- | --- |
|  |  |
| **Figure 5. Experiment 2- Visual Masking-** **Indication of Recovery**. In order to evaluate between-group differences in extinction retention a recovery index (RI) was calculated by deducting the CS elicited threat response at the last four trial of extinction from the trials of the Test Phase. Univariate analysis confirmed a significant group difference on the RI, where the control group (M = .0608) showed clear indications of recovery, whereas no recovery was observed in the unaware (M = -.0546) and the aware groups (M = -.0467), yielding an effect size (partial η2) of 0.42. |

  |  |

 |