The Role of Contingency Awareness and Unconditioned Stimulus Modality in Human Aversive Conditioning

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Yibo Li
aus Henan/China

Tübingen
2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 05.12.2025
Dekan: Prof. Dr. Thilo Stehle
1. Berichterstatter/-in: Prof. Dr. Martin Hautzinger
2. Berichterstatter/-in: apl. Prof. Dr. Boris Kotchoubey

Acknowledgements

I would like to express my heartfelt gratitude to everyone who supported me throughout the course of this work, conducted at the Institute of Medical Psychology and Behavioral Neurobiology at the Eberhard Karls University of Tübingen.

First and foremost, my sincere thanks go to Professor Dr. Boris Kotchoubey and Dr. Yuri G. Pavlov for proposing the topic of this research, for their continuous guidance, and for the time and expertise they generously shared. Their critical insights, thoughtful suggestions, and meticulous corrections greatly enriched the quality of this work.

I am also deeply grateful to my colleague Nick S. Menger, whose close involvement and valuable input were vital to every stage of the project. I would also like to thank Maria Zamojska-Turska for her kind assistance with data collection, which was essential to the success of this study.

Special thanks go to my wife, Rui Wang, whose love, companionship, and support sustained me every step of the way. I also thank my grandparents and parents for their unwavering support and belief in me. Their presence in my life made this journey possible.

Many thanks also to Professor Dr. Martin Hautzinger and Professor Dr. Boris Kotchoubey for reviewing this work.

Last but not least, I extend my appreciation to all participants and staff members at the Institute who contributed in various ways to the successful completion of this project.

Statement of Contributions

Section 2 of this thesis is based on the following publication:

Li, Y., Menger, N. S., Keil, A., Kosachenko, A. I., Mertens, G., Kotchoubey, B., & Pavlov, Y. G. (2025). **Contingency awareness and fear conditioning: A comprehensive examination of associated factors**. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001501

Authors contribution to the publication:

Yibo Li: conceptualization, investigation, visualization, methodology, data curation, project administration, formal analysis, writing-original draft, and writing-review and editing. **Nick S. Menger**: conceptualization, methodology, investigation, and writing-review and editing. **Andreas Keil**: conceptualization, methodology, and writing-review and editing. **Alexandra I. Kosachenko**: resources and writing-review and editing. **Gaëtan Mertens**: conceptualization, methodology, and writing-review and editing. **Boris Kotchoubey**: conceptualization, methodology, funding acquisition, and writing-review and editing. **Yuri G. Pavlov**: conceptualization, visualization, formal analysis, methodology, project administration, supervision, writing-original draft, and writing-review and editing. All authors contributed to the final manuscript and read and approved the final manuscript.

Abbreviations

US	Unconditioned stimulus/stimuli
UR	Unconditioned response
CS	Conditioned stimulus
CR	Conditioned response
SCR	Skin conductance response
FPS	Fear-potentiated startle
HR	Heart rate
EEG	Electroencephalography
ERP	Event-related potential
WM	Working memory
IUS	Intolerance of Uncertainty Scale
BIS	Behavioral Inhibition System
BAS	Behavioral Activation System
STAI	State-Trait Anxiety Inventory
NAPS	Nencki Affective Picture System
ITI	Intertrial interval
IMCs	Instructional Manipulation Checks
ECG	Electrocardiogram
ICA	Independent component analysis
LPP	Late positive potential
SPN	Stimulus-preceding negativity

Zusammenfassung

Aversive Konditionierung dient als zentrales Modell zur Untersuchung des menschlichen aversiven Lernens, wobei die Furchtkonditionierung dessen Kernform darstellt. Trotz umfangreicher Forschung bleiben in diesem Bereich zwei zentrale Fragen ungelöst. Die erste lautet, ob Kontingenzbewusstsein - definiert als das bewusste Erkennen der Beziehung zwischen dem konditionierten Stimulus (CS) und dem unkonditionierten Stimulus (US) - eine notwendige Bedingung für den Erwerb konditionierter Reaktionen (CRs) in der Furchtkonditionierung darstellt. Diese Frage hat wichtige Implikationen für Theorien des Furchtlernens und ist auch relevant für das Verständnis der Entwicklung von Angst- und furchtbezogenen Störungen sowie für die Gestaltung gezielter Interventionen. Die zweite Frage betrifft, wie die sensorische Modalität des US die CRs auf verschiedenen Reaktionsebenen beeinflusst, darunter die evaluative (subjektive), die autonome und die neuronale Domäne. Eine Klärung dieses Zusammenhangs kann das Verständnis des aversiven Lernens vertiefen und die Auswahl von US-Modalitäten bei der Untersuchung spezifischer CRs im experimentellen Design leiten.

Zur Beantwortung dieser Fragen wurden zwei komplementäre empirische Studien durchgeführt. Studie 1 nutzte ein groß angelegtes Online-Design mit 895 Teilnehmern, die 12 experimentellen Gruppen zugeordnet wurden. Sie untersuchte, wie prozedurale Parameter und individuelle Unterschiede, wie Persönlichkeitsmerkmale und kognitive Fähigkeiten, die Entstehung des Kontingenzbewusstseins in der Furchtkonditionierung beeinflussen. Zudem wurde getestet, ob Kontingenzbewusstsein notwendig ist für das Entstehen evaluativer CRs, gemessen anhand affektiver Bewertungen der CSs. Die Ergebnisse zeigten, dass mehrere prozedurale Faktoren - darunter Online-Furchtbewertungen, US-Erwartungsbewertungen, die Bearbeitung gleichzeitiger Aufgaben, die sensorische Modalität von CS und US sowie die bereitgestellten Instruktionen - das Auftreten von Kontingenzbewusstsein signifikant beeinflussten. Darüber hinaus zeigten nur Teilnehmer, die ein Bewusstsein für die CS-US-Beziehung aufwiesen, eine klare evaluative Differenzierung zwischen CS+ und CS-. Dieses Ergebnis deutet darauf hin, dass Kontingenzbewusstsein eine entscheidende Rolle bei der Bildung evaluativer CRs während des Furchtlernens spielt.

Aufbauend auf diesen Befunden wurde Studie 2 in einem kontrollierten Laborumfeld durchgeführt. Sie hatte das Ziel, die Rolle des Kontingenzbewusstseins weiter zu untersuchen und zu prüfen, wie unterschiedliche US-Modalitäten CRs in evaluativen, autonomen und neuronalen Systemen prägen. Die Teilnehmer wurden zufällig zugewiesen, entweder Elektroschocks, Luftstöße ins

Auge oder aversive Bilder als US zu erhalten. CRs wurden mittels affektiver Bewertungen, Hautleitungsreaktionen (SCRs), Herzfrequenz (HR) und elektroenzephalographischer (EEG) Aktivität gemessen. Die Ergebnisse zeigten, dass Elektroschocks die robustesten und konsistentesten CRs über alle Modalitäten hinweg auslösten, einschließlich höherer Furcht- und Erregungsbewertungen, niedrigerer Valenzbewertungen, erhöhter SCRs, HR-Verlangsamung, negativerer Stimulus-Preceding Negativity (SPN)-Amplituden, verstärkter Late Positive Potential (LPP)-Amplituden und stärkerer Alpha-Band-Powersuppression für CS+ im Vergleich zu CS-. Luftstöße führten zu verlässlicher HR-Verlangsamung und negativeren SPN-Amplituden für CS+ im Vergleich zu CS-, jedoch waren die evaluativen CRs schwach und es traten keine SCR-Effekte auf. Im Gegensatz dazu riefen aversive Bilder robuste evaluative CRs hervor, erzeugten jedoch keine nachweisbaren autonomen oder neuronalen CRs. Dabei zeigte sich, dass innerhalb der Bildgruppe das Kontingenzbewusstsein evaluative CRs modulierte, jedoch keinen Einfluss auf autonome Reaktionen zeigte. Insgesamt unterstreichen diese Befunde die zentrale Rolle der US-Modalität bei der Prägung der Stärke und des Reaktionsprofils aversiven Lernens. Zudem deuten sie darauf hin, dass Kontingenzbewusstsein in erster Linie mit evaluativen CRs verknüpft ist, während sein Einfluss auf autonome Systeme weniger eindeutig bleibt.

Zusammenfassend liefern diese Ergebnisse robuste Belege dafür, dass sowohl kognitive Faktoren wie das Kontingenzbewusstsein als auch stimulusbezogene Merkmale wie die US-Modalität zentrale Funktionen im menschlichen aversiven Lernen erfüllen. Kontingenzbewusstsein scheint eine entscheidende Rolle zu spielen für das Entstehen evaluativer CRs, wohingegen die sensorische Modalität des US sowohl die Stärke des Lernens als auch die spezifischen beteiligten Reaktionsebenen prägt, einschließlich evaluativer, autonomer und neuronaler Domänen. Darüber hinaus wurde das Auftreten von Kontingenzbewusstsein durch mehrere prozedurale Faktoren beeinflusst. Diese Befunde tragen zu einem differenzierteren Verständnis der Mechanismen des menschlichen aversiven Konditionierens bei und haben wichtige Implikationen für klinische Interventionen bei Angst- und furchtbezogenen Störungen.

Summary

Aversive conditioning serves as a key model for studying human aversive learning, with fear conditioning representing its central form. Despite extensive research, two core questions remain unresolved in this field. The first is whether contingency awareness, defined as the conscious recognition of the relationship between the conditioned stimulus (CS) and the unconditioned stimulus (US), constitutes a necessary condition for the acquisition of conditioned responses (CRs) in fear conditioning. This question has important implications for theories of fear learning and is also relevant to understanding the development of anxiety- and fear-related disorders as well as the design of targeted interventions. The second question concerns how the sensory modality of the US influences CRs across multiple response systems, including evaluative (subjective), autonomic, and neural domains. Clarifying this relationship can deepen understanding of aversive learning and inform the selection of US modalities when targeting specific CRs in experimental design.

To address these questions, two complementary empirical studies were conducted. Study 1 employed a large-scale online design with 895 participants assigned to 12 experimental groups. It examined how procedural parameters and individual differences, such as personality traits and cognitive functioning, influence the emergence of contingency awareness in fear conditioning. It also tested whether contingency awareness is necessary for the emergence of evaluative CRs, as measured by affective ratings of the CSs. The results showed that several procedural factors, including online fear ratings, US-expectancy ratings, engagement in concurrent tasks, the sensory modality of both the CS and the US, and the instructions provided, significantly influenced the emergence of contingency awareness. Moreover, only participants who demonstrated awareness of the CS-US relationship showed clear evaluative differentiation between CS+ and CS-. This finding suggests that contingency awareness plays a critical role in the formation of evaluative CRs during fear learning.

Building on these findings, Study 2 was conducted in a controlled laboratory setting to further examine the role of contingency awareness and to investigate how different US modalities shape CRs across evaluative, autonomic, and neural systems. Participants were randomly assigned to receive electric shocks, airpuffs to the eye, or aversive images as the US. CRs were measured using affective ratings, skin conductance responses (SCRs), heart rate (HR), and electroencephalographic (EEG) activity. The results showed that electric shocks elicited the most robust and consistent CRs across all systems, including higher fear and arousal ratings, lower valence ratings, elevated SCRs, HR deceleration, more negative stimulus-preceding negativity (SPN) amplitudes, enhanced late positive potential (LPP) amplitudes, and greater alpha-band power suppression for CS+ compared to CS-. Airpuffs evoked reliable HR deceleration and more negative SPN amplitudes for CS+ relative to CS-, but evaluative CRs were weak and no SCR effects emerged. In contrast, aversive images elicited robust evaluative CRs but did not produce detectable autonomic or neural CRs. Importantly, within the image group, contingency awareness modulated evaluative CRs but showed no impact on autonomic responses. Taken together, these findings underscore the pivotal role of US modality in shaping the strength and the response profile of aversive learning. They further suggest that contingency awareness is primarily linked to evaluative CRs, whereas its influence on autonomic systems remains less conclusive.

Overall, these findings provide robust evidence that both cognitive factors such as contingency awareness and stimulus-related features such as US modality play critical roles in human aversive learning. Contingency awareness appears especially important for the emergence of evaluative CRs, whereas the sensory modality of the US shapes both the strength of learning and the specific response systems involved, including evaluative, autonomic, and neural domains. In addition, the emergence of contingency awareness was influenced by several procedural factors. These findings contribute to a more nuanced understanding of the mechanisms underlying human aversive conditioning and carry important implications for clinical interventions targeting anxiety and fear-related disorders.

Table of Contents

1 General Introduction

1.1 Classical Conditioning

To adapt and thrive, humans must learn to recognize both meaningful and incidental patterns in the environment (Eelen, 2018). Classical conditioning is a key theoretical model for describing this type of learning process. Although American psychologist Edwin Twitmyer reported findings related to classical conditioning before Ivan Pavlov, a Russian physiologist, it was Pavlov's systematic and well-documented research that gained broader recognition (Pavlov, 1927; Rehman et al., 2025). His name subsequently became closely associated with the phenomenon, and the term "Pavlovian conditioning" remains widely used in both scientific discourse (Barrett et al., 2025; Harris, 2025; Lovibond & Westbrook, 2024) and everyday language.

Notably, Pavlov did not initially intend to study learning. While investigating the digestive processes of dogs, he noticed an unexpected pattern. The dogs began to salivate not only when food was presented, but also when they were exposed to cues that regularly preceded feeding, such as the sound of a cart. This observation prompted a series of experiments exploring the relationship between environmental stimuli and behavioral responses. These early studies exemplify appetitive conditioning, as they involved pairing neutral cues with a positive stimulus such as food, which naturally elicited a biological response.

In one experiment, Pavlov rang a bell just before offering food to the dogs. Initially, the bell had no effect on its own. However, after repeated pairings, the dogs began to associate the bell with the arrival of food. Eventually, they began secreting saliva upon hearing the bell, even when no food was presented (Pavlov, 1927). In this setup, the food is known as the unconditioned stimulus (US) since it naturally triggers salivation, referred to as the unconditioned response (UR). The bell initially functioned as a neutral stimulus, meaning it did not naturally elicit responses such as salivation. After being consistently paired with the food, the bell became a conditioned stimulus (CS) that could elicit salivation on its own, a response that was considered the conditioned response (CR). This experiment demonstrates that a neutral stimulus, after repeated pairings with a US, can ultimately elicit a CR. As such, this associative learning process represents the core of classical conditioning.

However, appetitive conditioning, as used by Pavlov, is often less feasible in adult human research due to methodological challenges. Food rewards involve variability in hunger, preferences, and dietary restrictions, and rapid satiation limits the number of comparable trials. Standardizing delivery across participants is also technically demanding. For these reasons, aversive conditioning paradigms, such as pairing stimuli with loud noises or mild electric shocks, are generally more efficient and reproducible.

1.2 Aversive Conditioning

Aversive conditioning, as one variant of classical conditioning, entails the consistent pairing of a neutral stimulus with an aversive US, for instance, an electric shock (Pavlov & Kotchoubey, 2019; Sheerin et al., 2025), loud noise (Abramson et al., 2024; Bacigalupo & Luck, 2018), noxious heat (Schlitt et al., 2022), aversive images (San Martín et al., 2023), or other unpleasant stimuli. Through this pairing, the neutral stimulus is transformed into a CS capable of producing CRs including negatively valenced affective responses, physiological arousal, expectancy and/or behavioral avoidance of US (Lonsdorf et al., 2017). These responses reflect defensive adaptations that support the detection of, preparation and response to potential threats.

Aversive conditioning plays a dual role in both research and clinical contexts. In experimental settings, it is widely used to model the pathological mechanisms and therapeutic processes underlying affective disorders such as anxiety and fear-related conditions (Lonsdorf et al., 2017). Clinically, it has also been applied to reduce various maladaptive behaviors in both children and adults (Strand, 2005).

1.3 Fear Conditioning

As a subtype of aversive conditioning, fear conditioning involves pairing a neutral stimulus with a threat-related US (Lonsdorf et al., 2017). This form of learning is considered evolutionarily adaptive, as it enables organisms to detect and respond to environmental threats. The first documented fear conditioning experiment in humans was conducted by Watson and Rayner in 1920 (Watson & Rayner, 1920). In this experiment, a child named Albert came to fear a white rat after its repeated pairing with a loud aversive noise. The learned fear subsequently extended to related objects, such as a rabbit and a fur coat, illustrating stimulus generalization. Although the experiment has been widely criticized for ethical reasons, it provided early empirical support for the idea that fear can be learned through associative processes.

Building on this foundational work, contemporary research typically employs differential fear conditioning paradigms (Lonsdorf et al., 2017). In such paradigms, a conditioned stimulus (CS+) is repeatedly linked to an aversive US and becomes a threat-related cue, whereas a similar stimulus (CS-) is never paired with the US and functions as a safety cue. A successfully conditioned fear response is typically characterized by heightened subjective, physiological, and behavioral reactivity to the CS+ relative to the CS-. That is, the discrimination between threat and safety cues represents a core feature of fear learning.

1.3.1 CS

Fear conditioning research commonly employs discrete exteroceptive CSs, with a strong emphasis on visual CSs such as geometric shapes (Bacigalupo & Luck, 2018; Yu et al., 2024), colored lights (Babel et al., 2018; Grillon et al., 2006), human faces (Ney, O'Donohue, et al., 2022; Reinhard et al., 2024), and images of animals (Ney, Nichols, et al., 2023). Beyond visual stimuli, research has also utilized other sensory modalities, including auditory tones (Bolaram et al., 2020), odors (Leer et al., 2011), and tactile sensations (Harvie et al., 2016) as CSs. Additionally, proprioceptive stimuli exemplified by arm movements (Meulders et al., 2015), and interoceptive stimuli including breathing loads (Pappens et al., 2013) or breathing in CO_2-enriched air (Acheson et al., 2007), have been adopted.

Although CSs are frequently selected for their emotional neutrality prior to learning, this is not always so. Several studies have used emotionally salient stimuli, including angry male faces (Raes et al., 2010; Raes & De Raedt, 2011), and fear-relevant animals like snakes and spiders (Katkin et al., 2001; Lipp et al., 2014), to investigate fear learning under more ecologically relevant conditions.

1.3.2 US

In fear conditioning studies, US are typically chosen for their ability to evoke defensive responses without prior learning. Painful stimuli, for example electro-tactile stimulation, are frequently applied (Lonsdorf et al., 2017; Mertens & Engelhard, 2020). In populations where pain is inappropriate, milder options like airblasts to the throat (Grillon et al., 2004) or loud sounds, including white noise (Abramson et al., 2024; Reinhard et al., 2024), tones (Neumann & Waters, 2006), and screams (Björkstrand et al., 2022; Lambert et al., 2021; Poplin et al., 2024), have been effective. Visual stimuli, such as negative images (San Martín et al., 2023) or aversive video clips (Pejic et al., 2013), have also been tested. Interoceptive stimuli, including breathing loads (Pappens et al., 2013), airflow occlusion (Pappens et al., 2012), inhalation of CO_2-enriched air (Meulders et al., 2010), balloon expansion (Ceunen et al., 2015), rectal expansion (Icenhour et al., 2015), and thermal stimulation (Jenewein et al., 2013), are often used in studies on bodily threat.

1.3.3 Phases

The fear conditioning paradigm is usually subdivided into several key phases: pre-conditioning, acquisition, generalisation, extinction, and return of fear, each representing a distinct process in the learning of fear responses. Some studies, however, include only two phases: acquisition and extinction.

Pre-conditioning phase. This is the first phase of the fear conditioning paradigm and primarily includes calibration of the US and habituation to the CSs. A key component is US calibration, as the aversiveness of the US influences both attention and learning. In lab-based studies using electric shocks, calibration typically follows a stepwise procedure, where participants rate increasing intensities until reaching an intensity perceived as unpleasant but not painful, or in some cases, significantly painful yet tolerable (Lonsdorf et al., 2017).

CS habituation involves presenting the CSs without pairing them with the US. This helps to establish a baseline, to reduce initial orienting responses, and to familiarise participants with the rating procedures (Lonsdorf et al., 2017).

Fear acquisition phase. During this phase, the CS+ undergoes repeated pairings with the aversive US, leading to the development of a conditioned fear response (Liu et al., 2025). A perceptually similar but unpaired stimulus, the CS-, is also presented and functions as a safety signal. Across repeated pairings, individuals acquire the ability to distinguish between the threat cue (CS+) and the safety cue (CS-).

Fear generalisation phase. This phase examines whether the conditioned fear response extends to novel stimuli sharing perceptual or conceptual similarity with the CS+ (Dou et al., 2023). Generalisation is commonly observed as a response gradient, where fear reactions are stronger to stimuli that bear closer similarity to the original CS+.

Fear extinction phase. During this phase, the CS+ is repeatedly presented in the absence of the US, resulting in a gradual decline of the conditioned fear response (Rashidi et al., 2025). This process reflects new learning that the previously threatening stimulus no longer signals danger. The CS- continues to be presented and retains its function as a safety signal.

Return of fear phase. Following extinction, conditioned fear responses may re-emerge under specific conditions, reflecting the persistence of the original fear memory. Several phenomena illustrate this return. Spontaneous recovery refers to the reappearance of the fear response after a delay without further exposure to the CS or US (Gershman & Hartley, 2015). Renewal occurs when a change in context between extinction and testing leads to the return of fear. For example, fear may be acquired in one context, extinguished in another, and may return upon presentation of the CS in either the acquisition context or a new context (Landkroon et al., 2019). Reinstatement involves the return of the fear response after an unexpected exposure to the US following extinction, even without any further CS-US pairing (Felmingham et al., 2021). These effects demonstrate that extinction results in new, context-dependent learning rather than the erasure of the original fear association.

1.3.4 CRs

Fear conditioning responses can be assessed across multiple levels, typically classified into three domains: (1) physiological and neurobiological measures, (2) self-report measures, and (3) behavioral indicators. These domains reflect different facets of fear learning and are often used in combination to provide a comprehensive assessment.

Physiological and neurobiological measures. These measures provide objective indices of conditioned fear by capturing autonomic and neural responses associated with threat processing and learning.

Electrodermal activity

Electrodermal activity is one of the most widely used physiological indicators and includes two components: skin conductance responses (SCRs) and skin conductance level. SCR refers to phasic, stimulus-evoked increases in skin conductance and is typically greater for CS+ than CS- in cue conditioning paradigms (Lykken & Venables, 1971; Mancinelli et al., 2024). Skin conductance level, in contrast, reflects the average level of conductance over a longer period while excluding phasic fluctuations, and is commonly used in context conditioning to differentiate between threat-associated and safe environments (Lykken & Venables, 1971; Neueder et al., 2019).

Fear-potentiated startle

Fear-potentiated startle (FPS) is a widely used index of defensive reactivity, referring to the amplification of the startle reflex in the presence of threat-related cues. This reflex is typically measured via the eyeblink response, using electromyography to record activity in the orbicularis oculi muscle (Blumenthal, 2015; Lonsdorf et al., 2017). To elicit the startle response, brief aversive stimuli such as bursts of white noise or airpuffs are used as startle probes. These probes are presented either during or between trials involving CSs, allowing researchers to assess how learned threat associations influence reflexive responses (Lissek, Baas, et al., 2005; Lissek, Powers, et al., 2005). The amplitude of the startle response is typically greater during CS+ trials compared to CS- trials, reflecting heightened defensive responding to perceived threat.

Heart rate and pupillary responses

Conditioned heart rate (HR) responses are a widely used index of fear conditioning, characterized by either HR deceleration, which reflects an orienting response, or HR acceleration, which indicates a defensive reaction. These responses vary across individuals and are influenced by factors such as stimulus type and intensity (Lonsdorf et al., 2017). Pupillary responses represent a fast and sensitive measure of conditioned fear, reflecting psychological arousal and assessable via eye-tracking or pupillometry. In contrast to SCRs, pupil dilation occurs rapidly following stimulus onset and can capture both continuous and stimulus-evoked changes during fear conditioning (Lonsdorf et al., 2017).

Electroencephalography and magnetoencephalography

Electroencephalography (EEG) and magnetoencephalography offer high temporal resolution for investigating how sensory systems respond to CS (Lonsdorf et al., 2017). They capture the electric or magnetic fields primarily generated by large-scale synchronized postsynaptic activity in the apical dendrites of cortical pyramidal neurons (Olejniczak, 2006), typically reflected in event-related potentials (ERPs) or event-related fields (Stolarova et al., 2006). In addition to transient stimulus-locked responses, EEG and MEG can capture ongoing brain rhythms across delta, theta, alpha, beta, and gamma bands during human fear conditioning (Trenado et al., 2018). These rhythms are thought to reflect coordinated neural activity involved in perception and attention. As the signals measured by EEG and MEG arise predominantly from cortical sources, both techniques are limited to detecting activity within cortical networks and are relatively insensitive to subcortical regions. Despite their high temporal resolution, their spatial resolution is limited (Miskovic & Keil, 2012). Furthermore, a large number of trials per condition is often required to achieve an adequate signal-to-noise ratio (Sperl et al., 2016), which can be challenging in standard fear conditioning paradigms.

Functional magnetic resonance imaging

Functional magnetic resonance imaging provides high spatial resolution and has been extensively used to identify key brain regions involved in fear learning, including the amygdala, hippocampus, and medial prefrontal cortex (Biggs et al., 2020; Dieterich & Brandt, 2024; Fullana et al., 2016; Kampa et al., 2024). In addition to identifying activation in specific regions, advances in analytical approaches such as multivariate pattern analysis have enabled researchers to examine distributed patterns of neural activity associated with fear conditioning (Hennings et al., 2022; Levine et al., 2021; Staib et al., 2020; Staib & Bach, 2018). However, the low temporal resolution of fMRI limits its ability to capture the rapid dynamics of fear learning (Kim et al., 1997).

Self-report measures. Self-report measures in fear conditioning primarily assess two domains: contingency awareness, which reflects explicit knowledge of the CS-US relationship, and affective responses, which capture the affective impact of the CS.

Contingency awareness

Contingency awareness refers to participants' explicit understanding of which CS is followed by the US, representing higher-order cognitive processing during fear learning (Labrenz et al., 2015). It has traditionally been assessed using post-experimental questionnaires. In such assessments, participants are asked whether the US followed the CS+, the CS-, occurred unpredictably, or whether they could not tell. In addition to choosing one of these options, participants are also asked to rate how confident they are in their answer. Based on these responses, participants who correctly identify the CS+ as being followed by the US and report a moderate to high level of confidence are classified as being aware of the contingency (Dawson et al., 2007; Dawson & Biferno, 1973; Lovibond et al., 2011; Singh et al., 2013).

More recent studies have refined this method by asking participants whether they noticed a specific CS-US pairing and, if so, having them identify the paired CS from a set of options. Accurate identification is taken as evidence of contingency awareness (Constantinou et al., 2021; McGregor et al., 2021, 2022). However, post-experimental measures may underestimate awareness due to forgetting or interference, especially when administered after extinction (Lovibond & Shanks, 2002).

In addition to retrospective assessments, contingency awareness can also be inferred from online (trial-by-trial) US expectancy ratings. These ratings often use dichotomous responses (e.g., expected/not expected), Likert scales, or visual analog scales (Boddez et al., 2013). In some previous studies, the ratings ranged from 0 to 100, where 0 indicated certainty that the US would not occur, 50 indicated uncertainty, and 100 indicated certainty that the US would occur. A relatively strict operationalisation defines awareness as five consecutive trials in which the US expectancy rating exceeds 75 for CS+ trials and remains below 25 for CS- trials, with at least two of each trial type included in the sliding window (Schultz & Helmstetter, 2010; Sevenster et al., 2014; Singh et al., 2013). Other studies adopt the "dual contingency" criterion, which required subjects to express awareness of two contingencies, i.e., the positive expectancy of the US during CS+ and negative expectancy of the US during CS- for three consecutive pairs of trials (Biferno & Dawson, 1977; Dawson & Biferno, 1973; Marinkovic et al., 1989).

In some studies, expectancy ratings are not collected online during acquisition, but instead assessed retrospectively after the acquisition phase. Here, contingency awareness is inferred by subtracting the expectancy score for CS- from that for CS+, with any positive difference interpreted as awareness (Raes et al., 2009).

Importantly, there is no universally accepted standard for defining contingency awareness based on expectancy ratings. Moreover, it has been suggested that the act of employing trial-by-trial expectancy ratings may itself increase contingency awareness by directing attention to the CS-US relationship (Baeyens et al., 1990; Lonsdorf et al., 2017).

Overall, careful selection of measurement methods is required, and the use of multiple complementary measures within a single study is recommended to enhance reliability and minimize bias.

CS affective ratings

In addition to contingency awareness, participants' affective responses to CSs are commonly evaluated. These assessments capture emotional dimensions such as valence (e.g., "How pleasant is the feeling this picture evokes?"), arousal (e.g., "How intense or arousing is the feeling?"), fear or anxiety (e.g., "How anxious does this shape make you feel?"), liking, and distress (Lonsdorf et al., 2017). Affective ratings are typically collected using Likert scales, visual analog scales, or the Self-Assessment Manikin.

As with expectancy ratings, affective ratings may be collected online (i.e., trial-by-trial), intermittently (e.g., after stimulus blocks), or retrospectively (e.g., after the acquisition phase). While these measures provide valuable insight into the subjective experience of fear learning, their timing must be carefully considered, as in-task assessments may interfere with ongoing physiological or cognitive processes (Lonsdorf et al., 2017).

Behavioral indicators. Behavioral indicators provide a complementary means of assessing conditioned fear, although they are used less frequently than physiological, neurobiological, or self-report measures. In animal studies, overt defensive behaviors such as flight or freezing can serve as direct markers of conditioned fear (Furuyama et al., 2023; Trott et al., 2022). In human research, however, eliciting such reactions is often impractical or ethically inappropriate, as it would require creating intense, real threats that may result in significant psychological or physical discomfort to participants. Nevertheless, related but subtler defensive responses can be assessed in the laboratory, one example being body posture freezing (van Ast et al., 2022). Another commonly used behavioral measure is reaction time (Dirikx et al., 2007; Ney, FitzSimons-Reilly, et al., 2023).

Postural freezing

Postural freezing is typically defined as a reduction in postural sway while standing, as measured by changes in center-of-pressure excursions in the anterior-posterior and mediolateral directions (Roelofs et al., 2010). This measure is derived from changes in the center of pressure, with lower sway values indicating reduced body mobility and reflecting greater motor inhibition in the presence of threat-related cues (van Ast et al., 2022).

Reaction time

Reaction time is particularly suitable for online experiments where physiological recordings are difficult to obtain. For example, in an online study, participants viewed words from two semantic categories, with one designated as CS+ and the other as CS-. Each CS was displayed for 6 seconds. During acquisition, CS+ words were followed on 75% of trials by an aversive US consisting of a 3-second loud scraping noise paired with an image of fingernails on a chalkboard, whereas CS- words were never followed by the US. On half of all trials, a neutral 800 Hz tone was presented 500 ms to 2.5 s after CS onset, prompting a spacebar response. Reaction times were significantly slower for CS+ than CS- trials, reflecting enhanced attentional engagement or cognitive interference triggered by threat-related cues, which in turn slowed responses to the auditory probe (Ney, FitzSimons-Reilly, et al., 2023).

1.4 The Role of Contingency Awareness in Fear Conditioning

Whether conditioning can occur in the absence of contingency awareness remains a central and debated issue in associative learning research, with implications not only for theoretical models of fear learning but also for clinical practice concerning anxiety- and fear-associated disorders.

Lovibond and Shanks (2002) reviewed a broad range of conditioning domains beyond fear conditioning, including aversive conditioning, eyeblink conditioning, and evaluative conditioning, to examine the association between contingency awareness and CRs. Across these paradigms, CRs were generally observed only when participants showed knowledge of the CS-US contingency. Although a small number of studies reported conditioning without awareness, these were typically marked by methodological shortcomings such as inadequate assessments of awareness or insufficient controls for alternative explanations. Based on this pattern of evidence, the authors concluded that awareness is a prerequisite for conditioning, though not sufficient on its own. They proposed that conditioning is governed by a propositional learning system reliant on conscious inference, rather than by an automatic associative mechanism operating independently of awareness (Lovibond & Shanks, 2002).

Building on this framework, Mertens and Engelhard (2020) performed a systematic review and meta-analysis on research investigating fear conditioning under conditions intended to obscure the CSs or their contingency with the US. Their analysis included 41 studies that measured both conditioning outcomes and awareness. An evaluation of methodological quality revealed that many studies suffered from limitations such as ineffective masking procedures, imprecise awareness measurements, trial-order effects, and analytic flexibility. Notably, greater methodological rigor was associated with a lower likelihood of supporting conditioning without awareness. Their meta-analytic moderation analyses failed to identify any reliable experimental settings that permitted fear conditioning to be demonstrated without awareness. Additionally, funnel plot irregularities and p-curve evaluations indicated the presence of publication bias. Drawing on these results, the authors argued that there is insufficient evidence to support fear conditioning without contingency awareness (Mertens & Engelhard, 2020). Taken together, these reviews converge on the conclusion that contingency awareness constitutes a prerequisite for acquiring CRs.

However, despite these influential reviews, there remains a body of conflicting evidence that cannot be readily dismissed. Several studies have documented observations of fear conditioning when explicit contingency awareness was absent (Jovanovic et al., 2006; Knight et al., 2009; Labrenz et al., 2015; Raio et al., 2012; Schultz & Helmstetter, 2010; Steinberg et al., 2012; Wong et al., 2004), although many of these may be subject to the methodological limitations noted above. Nevertheless, the lack of convincing data on unaware fear learning in humans should not be taken as definitive proof of its nonexistence. Moreover, the assertion that Pavlovian conditioning is fully governed by propositional learning is difficult to reconcile with its robust demonstration in non-human organisms such as worms and insects. The findings underscore that the contribution of awareness to fear conditioning continues to be debated, highlighting the need for further investigation into whether and under what conditions learning can emerge in the absence of conscious awareness.

1.5 Outstanding Questions and Objectives

Although fear conditioning has been extensively investigated, the contribution of contingency awareness remains contested. Some evidence points to explicit awareness of the CS-US contingency as a prerequisite for the development of conditioned responses, whereas other studies report that learning can emerge without conscious knowledge. This divergence underscores the necessity of further empirical research to clarify the link between contingency awareness and fear conditioning.

A closely related issue concerns the factors that influence the development of contingency awareness. However, few studies have systematically examined how either experimental parameters or individual differences such as trait anxiety, working memory (WM) capacity, and attentional control contribute to awareness acquisition. Clarifying the relationship between contingency awareness and fear learning, as well as identifying the factors that facilitate awareness acquisition, is essential for understanding how individuals develop fear of novel stimuli and for informing more effective interventions for anxiety and fear-related disorders.

Another key question concerns how the US properties influence the acquisition of CRs. According to the Rescorla-Wagner model, the US affects both the rate of learning and the magnitude of CRs (Rescorla & Wagner, 1972). Although prior research has employed a variety of US modalities in aversive conditioning, such as electric shock, loud noise, airpuff, and aversive images, direct comparisons across these modalities remain limited (Glenn, Lieberman, et al., 2012; Neumann & Waters, 2006; Ney, Nichols, et al., 2023; Sperl et al., 2016). Moreover, most studies examining the effects of different US modalities have focused on a narrow set of outcome measures, such as self-reported affect or autonomic responses, with few incorporating neurophysiological indices. Consequently, the extent to which US modality shapes CRs across psychological, physiological, and neurobiological domains remains unclear.

To address these issues, I conducted two empirical studies that are complementary in focus and methodologically distinct.

Study 1 is a large-scale online study involving approximately 900 participants assigned to 12 experimental groups. The first objective is to examine the relationship between contingency awareness and fear learning, using affective ratings as the primary measure of CRs. The second objective is to examine factors influencing the acquisition of contingency awareness, including both experimental manipulations and individual difference variables such as demographic characteristics, anxiety-related traits, and cognitive abilities. This study provides an initial framework for understanding how conscious awareness of CS-US contingencies emerges and relates to fear-related evaluative responses across a diverse sample and a range of experimental paradigms.

Study 2 is a laboratory-based study involving 115 participants, with data collected across multiple measures, including affective ratings, physiological indicators (SCR, HR), and a neurophysiological index (EEG). One objective was to explore the association between contingency awareness and CRs by collecting both subjective affective ratings and physiological measures (SCR, HR), thereby testing the robustness of the findings from Study 1 and further deepening the understanding of the relationship between these two constructs. Another objective was to compare the effects of different US modalities (electric shocks, airpuffs, aversive images) on CRs across subjective, physiological, and neurophysiological systems. By addressing these aims, Study 2 extends the findings of Study 1 and provides new insights into how US modality shapes aversive learning across multiple levels of analysis.

Together, these two studies examine how contingency awareness and US modality influence CRs and thereby deepen understanding of the factors that shape aversive learning across psychological, physiological, and neurobiological responses.

2 Study 1: Contingency Awareness and Fear Conditioning: a Comprehensive Examination of Associated Factors

2.1 Introduction

Identification of threats and organization of appropriate responses to reduce or avoid harm is critical for survival. One of the mechanisms subserving this function is associative fear learning (Lonsdorf et al., 2017). This is typically studied in the fear conditioning paradigm. In this paradigm, an initially neutral stimulus, e.g., the ringing of a bell (conditioned stimulus, CS+), is repeatedly paired with an aversive stimulus (unconditioned stimulus, US) such as a loud noise, while another neutral stimulus, e.g., a whistle (CS-) is never paired with the US. After several pairings, CS+ begins to trigger the conditioned response, which entails a subjective, physiological and/or behavioral reaction that is preparatory for the onset of the US. One of the central debates in fear conditioning revolves around contingency awareness, which refers to the ability to consciously recognize the association between CS and US (Lovibond & Shanks, 2002). Although some studies suggest that fear acquisition learning can occur without contingency awareness (Jovanovic et al., 2006; Knight et al., 2003; Raio et al., 2012; Schultz & Helmstetter, 2010; Steinberg et al., 2012), a meta-analysis including thirty original studies found no convincing evidence to support this claim (Mertens & Engelhard, 2020). However, previous research is primarily based on atypical fear conditioning tasks (e.g., involving masking of the CS) or small sample sizes, making the results inconclusive. Clarifying the relationship between awareness and fear conditioning[1] is crucial for understanding the conditions under which individuals learn to fear new stimuli, and can help to understand the aetiology and treatment of fear-related disorders.

A key, yet often overlooked, question in this debate is how many participants actually become aware of the CS-US association in typical conditioning experiments. If a substantial subset remains unaware, we need to determine whether these individuals truly acquire fear without awareness, and which methodological factors in standard procedures inadvertently prevent or

---

[1] Our operational definition of “fear conditioning” is the development of fear, orienting or defensive responses specifically resulting from repeated CS–US pairings. These responses can be observed both in subjective expressions (e.g., “I am fearful”) and in characteristic physiological correlates. In the present study, this operational definition does not include any cognitive changes induced by CS–US pairings, such as contingency awareness or explicit expectations of the US in the presence of the CS.

obscure awareness. Depending on the experimental design, ~15-30% of participants will never acquire contingency awareness (Berg et al., 2023; Mertens & Engelhard, 2020). Why do people fail in such seemingly simple tasks? Being able to focus all attention on the task seems to be an important factor - some studies have successfully prevented participants from gaining contingency awareness by using distracting simultaneous tasks during conditioning (Clark & Squire, 1998; Dawson et al., 1986; Manns et al., 2001; Weidemann et al., 2016). Variations in instructions (fully or partially informing participants about the contingencies vs. not informing them) have also been shown to significantly impact awareness levels (Dawson & Biferno, 1973; Mertens et al., 2021). However, many potentially relevant procedural parameters have never been systematically investigated. For example, it is plausible that trace conditioning that includes some temporal gap between CS and US would decrease the proportion of contingency aware participants as compared with delay conditioning where the US occurs while the CS is still presented or immediately at its offset. However, no studies have explored this possibility. Other theoretically possible but never studied factors are the rate of reinforcement, the nature of the US (e.g., loud noise or aversive pictures), and modality of the CS.

Additionally, inter-individual differences in personality traits and cognitive abilities may modulate the probability of becoming aware of contingencies. The effect of elevated trait anxiety - a commonly studied variable affecting acquisition of fear - was shown to increase (Rehbein et al., 2015), decrease (Chan & Lovibond, 1996), or not affect (Berg et al., 2023; Sevenster et al., 2014) contingency awareness. Central cognitive abilities such as working memory capacity are likely also relevant for contingency awareness, such as for supporting the ability to maintain the CS in memory if the trace conditioning procedure is involved. The relationship of contingency awareness and working memory was tested and found to be positive in a small group of participants (Cosand et al., 2008). Another contributing concept related to cognitive control is mind wandering, which refers to the diversion of attention from the current task and situation (Schooler et al., 2011). We hypothesise that there is a negative correlation between proneness to mind wandering and awareness acquisition. Overall, exploring the modulating effects of inter-individual differences on awareness acquisition is essential to understanding the phenomenon.

The primary limitation of studying contingency awareness in a laboratory setting is the inherent imbalance in participant groups, where, in a typical scenario, about 85% become aware, and 15% do not. Therefore, many previous studies on fear conditioning in relation to contingency awareness suffer from limited statistical power to detect small and medium size effects. This limitation can be mitigated by conducting studies online. The emergence of recruiting platforms such as Prolific and Javascript-based experimental platforms such as Pavlovia have rapidly expanded the recruitment of large and diverse global populations, which can enhance statistical power and also increase the external validity of studies. Recently, several researchers have begun to conduct online fear conditioning studies demonstrating the feasibility of such an approach (Björkstrand et al., 2022; Lambert et al., 2021; Plog et al., 2023; Purves et al., 2023; Stegmann et al., 2021).

Here, using a large online sample, we aimed to systematically study the role of contingency awareness and possible procedural and inter-individual moderators thereof in human fear conditioning to inform the research community about this ongoing theoretical debate. To that end, we conducted twelve experiments, each employing a variation of a prototypical fear conditioning task. These tasks were not substantially modified in ways that would obscure learning (such as masking) or involve major methodological deviations from standard fear conditioning procedures typically employed in a lab setting (Bach et al., 2023). Collectively, our experiments were designed to address three key research questions.

First, to answer the question which experimental parameters affect the contingency awareness rates, we tested whether parameters such as CS-US temporal contiguity (i.e., delay vs trace conditioning), the level of US aversiveness, reinforcement rate, and other parameters were able to modulate contingency awareness rates. The experimental parameters and associated hypotheses are visually summarized in **Figure 1a**.

Second, to test whether fear learning without awareness is possible, we compared indicators of learning (as expressed in affective ratings of CSs[2]) between contingency aware and unaware individuals. We hypothesized that participants who are contingency unaware would not exhibit fear learning, as expressed in no CS+/CS- difference in the affective ratings.

Third, to test whether individual differences in personality traits affect the probability of becoming contingency-aware, we measured several traits previously associated with fear conditioning: trait anxiety, state anxiety, intolerance of uncertainty, the behavioural inhibition system (BIS) and the behavioural activation system (BAS) (Lonsdorf et al., 2017; Lonsdorf & Merz, 2017; Morriss et al., 2019). Furthermore, we measured working memory (WM) capacity to understand how the ability to maintain information over short periods of time is related to the ability to make associations between events. We also used WM as a measure of general cognitive ability, given its association

---

[2] Changes in the valence of the CS after the CS is paired with an emotional outcome are also known as "evaluative conditioning."

with various cognitive abilities such as fluid intelligence (Unsworth et al., 2014), language comprehension (Daneman & Merikle, 1996), and mathematical reasoning (Raghubar et al., 2010). We expected a positive correlation between WM capacity and contingency awareness attainment. We also examined susceptibility to mind wandering as an indicator of attention control, hypothesizing a negative correlation with contingency awareness acquisition. Lastly, we predicted a negative association between age and the acquisition of contingency awareness, due to older age being linked to declining cognitive capacity and sustained attention (Burns & Zaudig, 2002).

A comprehensive list of the hypotheses corresponding to the research questions outlined above is presented below.

Hypothesis 1: Participants who are not consciously aware of the relationship between the CSs and the US will not exhibit fear learning, as measured by changes in the affective ratings of the CSs: (1.1) arousal, (1.2) fear, and (1.3) valence.

Hypothesis 2: The inclusion of intermittent online fear ratings (Group 2) will decrease the proportion of participants who develop contingency awareness as compared to a design without such ratings (i.e., interference effect) (Group 1).

Hypothesis 3: The inclusion of intermittent online US-expectancy ratings (Group 3) will decrease the proportion of participants who develop contingency awareness as compared to a design without such ratings (i.e., interference effect) (Group 1).

Hypothesis 4: A 55% reinforcement rate (Group 4) will reduce the proportion of participants who develop contingency awareness in comparison with a 100% reinforcement rate (Group 1).

Hypothesis 5: The use of a trace conditioning procedure (Group 5, in which the US is presented 2 seconds after the end of the CS+ presentation) will reduce the percentage of participants who develop contingency awareness, compared to a delay conditioning procedure (Group 1, in which the US is presented immediately after the end of the CS+ presentation).

Hypothesis 6:

6.1 The inclusion of a habituation phase before the acquisition (Group 6) will reduce the proportion of participants who develop contingency awareness, compared to a procedure without a habituation phase (Group 1).

6.2 The inclusion of a habituation phase before the acquisition (Group 6) will decrease average differential fear ratings in the contingency aware group, compared to a procedure without a habituation phase (Group 1).

Hypothesis 7: Compared to a procedure without a concurrent task (Group 1), a concurrent task (listening to a story) during the acquisition phase (Group 7) will reduce the proportion of participants who develop contingency awareness.

Hypothesis 8: Using auditory stimuli (simple tones) as CSs (Group 8) compared to using geometrical shapes (Group 1) will reduce the proportion of participants who develop contingency awareness.

Hypothesis 9: Using a loud noise as the US (Group 9) will increase the proportion of participants who develop contingency awareness, compared to using unpleasant pictures as the US (Group 1).

Hypothesis 10: The presentation of a loud noise 500 ms (US) before the end of the CS+ presentation (Group 10) will increase the proportion of participants who develop contingency awareness, compared to the presentation of a loud noise immediately after the end of the CS+ presentation (Group 9).

Hypothesis 11: Compared to not instructing participants on the CS-US contingencies before the experiment (Group 2), both (11.1) unspecific instructions (Group 11, e.g. “one of the shapes will always be followed by an unpleasant picture”) and (11.2) specific instructions (Group 12, e.g. “the circle will always be followed by an unpleasant picture and the diamond will not”) will increase the proportion of participants who gain contingency awareness, (11.3) whereas the specific instruction will have a stronger effect than the unspecific instruction.

Hypothesis 12:

12.1.1 Trait anxiety and 12.1.2 state anxiety will be positively associated with the acquisition of contingency awareness.

12.2 Intolerance of uncertainty will be positively associated with the acquisition of contingency awareness.

12.3 The activity of the behavioral inhibition system will be positively associated with the acquisition of contingency awareness.

12.4 The activity of the behavioral activation system will be negatively associated with the acquisition of contingency awareness.

Hypothesis 13: Age will be negatively associated with the acquisition of contingency awareness.

Hypothesis 14: The working memory capacity will be positively associated with the contingency awareness rate.

Hypothesis 15: The number of mind wandering episodes will be negatively associated with the contingency awareness rate.

Hypothesis 16: Concerning the participant's emotional response to the US; (16.1) the degree of negative affect (valence rating) will be positively associated with the acquisition of awareness, and (16.2) the degree of arousal will be positively associated with the acquisition of awareness.

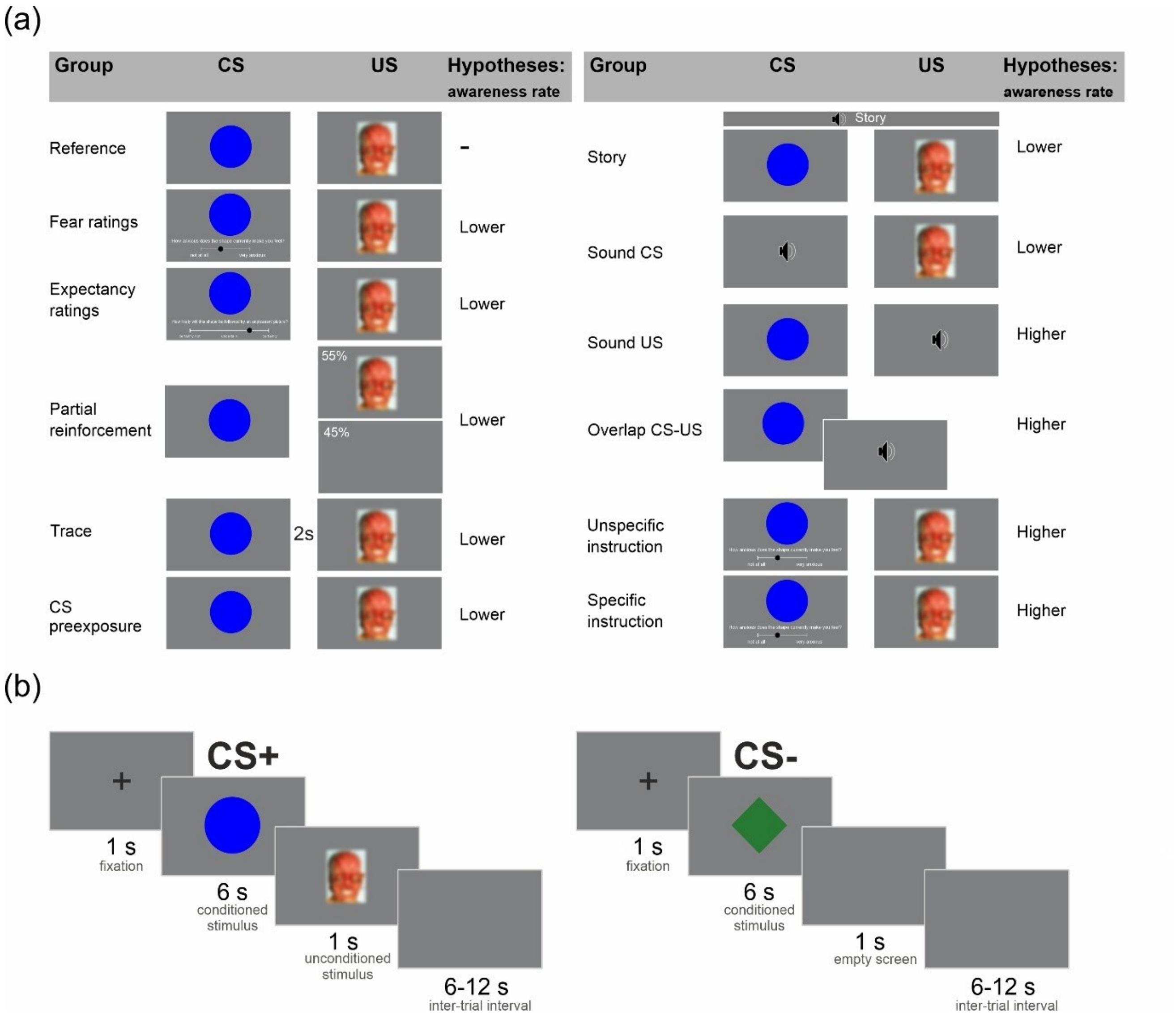


Figure 1. (a) Experimental parameters and hypotheses. "CS preexposure" group had a habituation phase before the fear acquisition phase during which no US were presented (identical to the extinction phase). In the groups "Unspecific instruction" and "'Specific instruction", the participants were partially or fully informed about CS-US contingency prior to the acquisition phase, respectively. For the "'Overlap CS-US" group, the "Sound US" group serves as the reference group. Both the 'Unspecific instruction' and "Specific instruction" groups use the "Fear ratings" group as their reference. For all other groups, the "Reference" group was used for comparison. (b) Trial examples used in the conditioning experiments. The US image of the mutilated face is blurred on this example image for ethical purposes and due to licensing agreements. The US was not blurred in the real experiments.

2.2 Methods

Experiments 1 and 2 were not preregistered and were used to formulate hypotheses and an analysis plan for the following experiments. The hypotheses and methods for experiments 3-12 were formally pre-registered (https://osf.io/g8dzs). Data are publicly available on OSF (https://osf.io/wvh8p).

2.2.1 Participants

The recruitment process for this study was carried out using Prolific (www.prolific.com). Only individuals fluent in English and within the age range of 18 to 60 were eligible to participate. The exclusion criteria for this study included: 1) any neurological or psychiatric disease, either current or in the past; 2) any acute disease in the past week that required medication; 3) pregnancy or lactation. Participants signed an electronic informed consent form before their participation in the study. A debriefing was provided to all participants at the end of the conditioning task.

The median participation time for the study was 43.54 minutes, with each participant receiving £6 as compensation. A total of 1048 individuals participated in the study (mean age = 32.90 years, age range = 18-60 years; 414 females, 634 males). In full compliance with the preregistration, we excluded 153 participants (14.6%) for low attention levels (refer to '**Attention checks**' below), leaving 895 participants (85.4%, mean age = 33.30 years, age range = 18-60 years; 356 females, 539 males) for statistical analyses. For groups 3-12, participants who failed attention checks were replaced until the target sample size of 75 was reached. Data collection for groups 1 and 2 occurred before preregistration, and those groups had slightly smaller final sample sizes due to a non-automated exclusion procedure at the time. All 895 participants were assigned into 12 groups with good homogeneity across groups, as detailed in **Appendix 6.1**. This study was approved by the Ethical Committee of the Medical Faculty of the University of Tübingen.

2.2.2 Procedure

Participants initially filled out a demographics questionnaire, followed by the personality questionnaires. Subsequently, they engaged in the working memory task, and finally, participated in the conditioning experiment.

Questionnaires. Information about sex, age, educational background, native language, drug use (including smoking, alcohol, coffee, and any other substances consumed within the past two hours), socioeconomic status, and handedness was collected from the participants. To assess personality traits, the 27-item Intolerance of Uncertainty Scale (IUS) (Freeston et al., 1994), the 20-item Behavioral Inhibition System/Behavioral Activation System Scales (BIS/BAS) (Carver & White, 1994), and the 40-item State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1983) were administered.

Working memory task. All participants completed 100 trials of a change localization task, an experimental paradigm used to assess visual working memory capacity (Zhao et al., 2023). In each trial, participants were shown a sample array of six colored squares (0.4 pixels × 0.4 pixels) for 400 milliseconds. After a 1000-millisecond blank screen, they were presented with a test array identical to the sample array, except that one square changed to a color not previously presented. The colors were selected randomly without replacement from a set of nine, with the following red, green, and blue values: [[1, -1, -1], [-1, 1, -1], [-1, -1, 1], [1, 1, -1], [-1, 1, 1], [1, -1, 1], [1, 0.2941, -1], [-1, -1, -1], [1, 1, 1]]. Participants had to identify and click on the square in the test array that had changed its color. The test array remained visible until a response was made. Trials were separated by a 1000-millisecond inter-trial interval. The percentage of correct answers served as an indicator of working memory capacity. A visual illustration of the task is shown in **Figure 2**, and a link to the working memory (WM) task is provided in **Appendix 6.2**.

Throughout the experiment, participants were pseudorandomly asked every 13-18 trials, for a total of six times, to describe their current conscious experience using six response options: 1) I was completely focused on the current task; 2) I was thinking about my performance on the task; 3) I was distracted by sights, sounds, or physical sensations; 4) I was intentionally thinking about things unrelated to the task; 5) I was unintentionally thinking about things unrelated to the task; and 6) My mind was blank. The responses where participants selected options 3 to 6 indicated a mind-wandering episode. The total number of episodes was used in the statistical analyses.

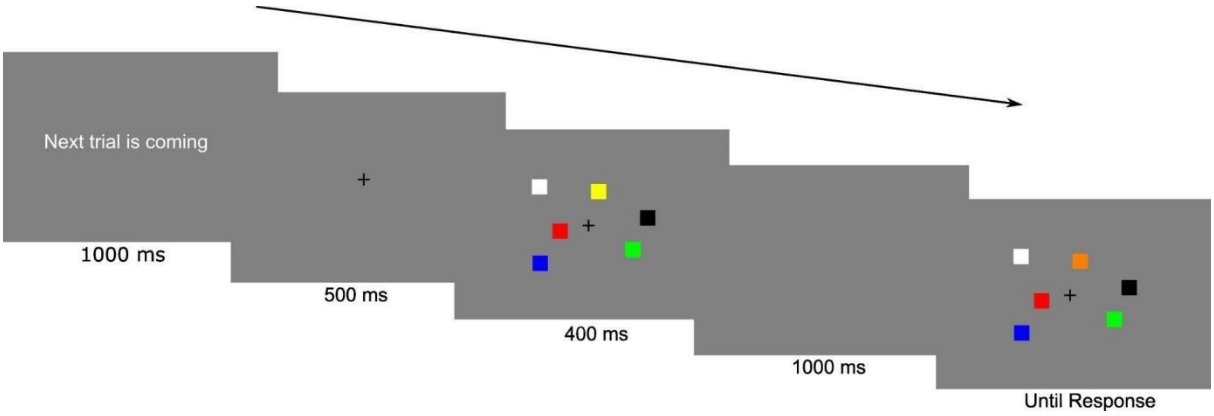


Figure 2. Sequence of events in a trial in the change localization task.

Sound calibration. In Group 7 (Story), to address variations in system volume among participants, we implemented the following calibration procedure. This procedure has been shown to be feasible and effective in a previous study (Berg et al., 2023). A 1-second 1200 Hz sound served as the test stimulus. Participants were exposed to this sound at two different volume levels before commencing the conditioning paradigm. Given that the decibel scale measures sound volume relatively, we used Audacity 3.3.2 to create two identical sounds except for a consistent 60 dB difference in loudness between them. The 60-dB difference remained constant regardless of the actual volume at which the sounds were played. During the calibration phase, participants first heard a very soft 1200 Hz sound and adjusted its volume using on-screen buttons until it was barely audible. If calibrated correctly by the participants, the louder sound was played at a level 60 dB sound pressure level above the hearing threshold. Using Audacity, we normalised the average volume of the story to that of the louder sound. This method ensured that, although the story was not played during the calibration phase, it would be presented at the intended volume of approximately 60 dB when the actual conditioning paradigm began. The calibration phase for Group 8 (Sound CS) was identical to the calibration phase for Group 7.

The calibration procedure in Group 9 (Sound US) and Group 10 (Overlap CS-US) closely resembled that of Group 7, with the main difference being the volume difference between the two test sounds, which was 85 dB instead of 60 dB. After adjusting the volume for the first test tone, participants encountered the second test sound (approximately 85 dB). Participants were then asked to evaluate whether this volume level was acceptable or painfully loud. If they found it to be intolerable, they could reduce the volume by clicking a button on the screen, which would decrease the loudness by 5 dB. Using Audacity, the volume of US was set to match the louder of the two test sounds. It should be noted that before deploying the paradigm, we conducted numerous tests using this calibration procedure on various computer devices and measured the US volume using a precise sound meter. In reality, the US volume was closer to 90 dB than to 85 dB. This is because when participants reported that they could barely hear the softer of the two 1200 Hz sounds, the softer sound was not at 0 dB.

Conditioning. There were a total of 12 groups, each differing from others in some aspects of the experimental paradigm. The selection of these aspects is the result of a series of meetings with six conditioning experts aiming to reach consensus on the conditions most important from a theoretical point of view. This was done based on each of the experts creating a rank-order of the most important procedural variations and subsequent consensus through discussion. A high degree of consensus was reached for the 12 procedural variations (groups) that should have the strongest effects on contingency awareness. Links to all conditioning tasks for Groups 1 to 12 are provided in **Appendix 6.2**.

Group 1: Reference group (Reference)

Circular and diamond-shaped images were used as the CS in the experiment. One of the images was green and the other was blue, with the allocation of the colors being randomized between participants. For each participant, one image was randomly chosen as CS+, whereas the other was the CS-. Thus, CS+ and CS- differed in two features (shape and color).

Pictures served as unconditioned stimuli (US). Ten highly unpleasant pictures from the Nencki Affective Picture System (NAPS) (Marchewka et al., 2014) were chosen on the basis of their normative ratings on fearfulness and disgust (detailed information in **Appendix 6.3**). The set included images of open wounds, disfigured faces, and mutilated human bodies. We used the method of matching fear and disgust values to divide the initial set of ten pictures into two similar groups of five. For each participant, one group of pictures was used randomly as the US during the acquisition phase, while the remaining five were used only in the US memory task. In PsychoPy, we utilized height units to scale the sizes of images relative to the window height. The 'US' image had dimensions (width, height) of (0.5, 0.5), indicating that both its width and height were 50% of the window's height. Similarly, the 'CS' image size was set at (0.3, 0.3).

The experiment was subdivided into several phases. During the baseline phase, participants were asked to rate the CS for valence (“How pleasant is the feeling that this picture evokes in you?”; 0 = "Very unpleasant", 50 = "Neutral", 100 = "Very pleasant"), arousal ("How intense or arousing is the feeling that this picture evokes in you? "; 0 = "Not at all", 100 = "Very intense") and fear ("How anxious does this shape currently make you feel? "; 0 = "Not at all", 100 = "Very anxious"). Participants were first asked to rate a white square, which served as a training to use the rating system, and after that rate the circle and diamond. The circle and the diamond were presented in a random order.

Before the acquisition phase, the participants were instructed to look at the screen at all times, and that unpleasant pictures would sometimes be presented, but they did not receive explicit instructions about the CS-US contingency. During the acquisition phase, participants completed a total of 22 trials (11 per CS). Each trial lasted for 7 seconds. After a 1-second fixation cross, the CS was presented for 6 seconds, after which a CS+ was always followed by a 1-second presentation of a US (100% reinforcement rate), and the CS- was followed by a blank screen. Each of the five selected US pictures was presented twice, with one picture used three times. Trials were separated by an intertrial interval (ITI) filled with a blank screen for 6-12 seconds. No more than two stimuli belonging to the same condition were presented consecutively. The schematic diagram of the conditioning procedure is shown in **Figure 1b**.

After the acquisition phase, we assessed participants' awareness of the contingencies. Participants were asked, "Did you notice that an unpleasant picture was shown after seeing a specific colored shape?" Those who answered "Yes" were shown both CS side by side and asked to select which one was followed by unpleasant pictures. Next, they rated their confidence in their choice on a scale from 0 ("Not at all") to 100 ("Absolutely confident"). Afterwards, participants rated the likelihood (i.e., their subjective expectancy) of an unpleasant picture following each CS on a scale (0 = "Certainly not", 50 = "Uncertain", 100 = "Certainly") and then repeated the affective ratings (valence, arousal, and fear) for the two CSs in random order.

The extinction phase commenced after the acquisition phase, following the same general procedure. It consisted of 14 trials in total (7 per CS), with no unpleasant pictures presented anymore. The number of extinction trials was determined using a MATLAB implementation of the Rescorla-Wagner model (see **Appendix 6.4** for details).

Participants were not informed before the extinction phase that unpleasant pictures would not be shown anymore. Similar to the acquisition phase, after extinction, participants were asked to provide their expectancy ratings for unpleasant pictures following each CS, as well as provide affective ratings for the two shapes.

After the extinction phase of the task, the US memory task was administered. The participants were presented with the five pictures that had previously served as the US, and an additional set of five novel pictures. For each picture, participants had to answer (Yes/No) if they had encountered this picture previously during the experiment. Finally, participants were requested to complete valence and arousal ratings for the US.

Group 2: Online fear ratings (Fear ratings)

The procedural parameters were identical to those of Group 1, except for the presence of intermittent fear ratings during presentation of the CS. In the acquisition phase, participants were asked to rate how anxious the shape made them feel using a visual analogue scale (0 = "Not at all", 100 = "Very anxious") during the presentation of CS on trials 1, 2, 11, 12, 21, and 22. The scale was placed underneath the CS (see **Figure 1** for a visual representation). Each set of two ratings included one CS+ and one CS- trial in a random order. No ratings appeared in the other trials. In the extinction phase, similar ratings were shown during CS presentation on trials 1, 2, 13, and 14.

Group 3: Online US-expectancy ratings (Expectancy ratings)

The procedural parameters were identical to those of Group 2, except for the presence of intermittent US-expectancy ratings during presentation of the CS. Using the same scale as was used for collecting expectancy ratings after acquisition, participants rated the expectancy during the presentation of CS in the same trials as described above for Group 2.

Group 4: Partial reinforcement rate (Partial reinforcement)

The procedural parameters were identical to those of Group 1, except for a 55% (6 times instead of 100% - 11 times) reinforcement rate during the acquisition phase. The first CS+ trial was always reinforced, and the remaining five CS-US pairings were distributed randomly over the acquisition phase.

Group 5: Trace conditioning (Trace)

The procedural parameters were identical to those of Group 1, except that a 2-s blank interval was inserted between CS+ offset and US presentation.

Group 6: CS preexposure

The procedural parameters were identical to those of Group 1, except for a habituation phase before acquisition. The habituation phase consisted of 14 trials in total (7 per CS) and was identical to the extinction phase of the task.

Group 7: Concurrent listening to a story (Story)

The procedural parameters were similar to Group 1, with the exception of stricter controlled inter-trial intervals (ITIs) and the addition of a concurrent task (listening to a story) during the acquisition phase.

Before the experiment began, participants were asked to put on headphones and to not remove them during the experiment. Additionally, they were instructed to set the computer system's volume to 100%. The participants were informed that they would hear an audio story and, after the end of the story, they would be asked questions on its content. Throughout the acquisition phase, participants were instructed to both listen to the story and look at the screen.

In Group 1, the blank screen during ITIs varied randomly between 6 and 12 seconds. In Group 7, 22 random integers between 6 and 12 were first generated under the condition that the sum of the 22 numbers was 198, i.e., 22 multiplied by 9. These 22 numbers were then randomly assigned to the 22 blank screen sections in ITIs. The objective of this procedure was to maintain the duration of the acquisition phase at precisely 6 minutes and 14 seconds to match the duration of the audio story and the average duration of the acquisition phase in Group 1. Prior to entering the extinction phase, participants answered five questions related to the aforementioned story (see **Appendix 6.5** for the list of the questions and specifics on the story).

Group 8: Tones as the CSs (Sound CS)

The procedural parameters were identical to those of Group 1, except for using sounds as CSs. Two tones (6-second long) with frequencies of 1000 Hz and 1400 Hz were used as CSs.

Group 9: Loud noise as the US (Sound US)

The procedural parameters were identical to those of Group 1, except for using a loud noise as US instead of unpleasant pictures. A 1,000 Hz sine wave was embedded into a 1-second burst of white noise to serve as the US.

Group 10: Overlap CS-US

The procedural parameters were identical to those of Group 9, except for the presentation of the US occurring 500 ms before the end of the CS+ presentation.

Group 11: Unspecific instruction

The procedural parameters were identical to those of Group 2, except for using a modified instruction before the acquisition phase. Instead of only stating that there will be shapes and occasional potential pictures, the following phrase was added: "one of the shapes will always be followed by an unpleasant picture. Your task is to find out which one".

Group 12: Specific instruction

The procedural parameters were identical to those of Group 11, except for making the instruction even more specific: e.g., "the circle will always be followed by an unpleasant picture and the diamond will not".

2.2.3 Attention checks

Two Instructional Manipulation Checks (IMCs) were presented in all groups. The IMCs explicitly instructed participants to complete a task in a certain way, aiming to verify whether they paid attention to the instructions given during the study. The first IMC occurred in the questionnaire phase, where participants were instructed: "The following questions are very simple to answer. We present a range of verbs, and all you have to do is respond to every question by selecting -2, i.e., the response button all the way on the left of the scale." The presented verbs included "Laughing," "Viewing," "Inspecting," "Walking," "Kneeling," "Enjoying," "Riding," "Getting," and "Driving." Participants who selected "-2" for all items were considered to have passed this attention check. The second IMC was introduced during the pre-acquisition phase of the conditioning task. Participants were instructed: "On the next screen, you will see a question with a scale like earlier in the experiment. However, this time you should not answer the question, i.e., do not select anything on the scale. This is an attention check. The scale will appear for 6 seconds, and after it has disappeared, the experiment will automatically continue." A square then appeared on the screen with the prompt: "As instructed, please do not respond to this rating." Below the prompt, a response scale was displayed, ranging from "Very unpleasant" to "Very pleasant." Participants who refrained from making any selection were deemed to have passed this IMC. The US memory task also served as an attention check, but was not implemented in Groups 9 and 10, where the US was a loud noise instead of pictures.

For groups where we employed auditory stimuli, specifically groups 7 to 10, we added another IMC to identify participants who did not wear their headphones for the whole duration of the experiment. Prior to the extinction phase, participants received an auditory instruction "Please choose number nine", while an on-screen instruction stated “Please follow the instruction”. A set of numbers from 0 to 9 was displayed on the screen for selection.

We also gathered self-reported attention levels from the participants in all groups. In all groups except 9 and 10, participants responded to a single question: "How frequently did you look away from the screen during the presentation of unpleasant pictures? Please answer honestly. Your response will NOT affect your reward or cause any other consequences." with responses ranging from 0 ("Never") to 100 ("Always"). In Groups 9 and 10, participants were asked two questions: the first regarding the frequency of removing their headphones during the presentation of loud noises, measured using a slider from 0% to 100%; and the second concerning whether they lowered the volume on their PCs during the presentation of the noise, with the response options being "Yes" or "No".

On the basis of these tests, participants were excluded from the statistical analyses if they met any of the following criteria: a. Non-compliance with instructions in two main IMCs (applicable to all groups); b. Attaining an accuracy score of less than 80% on the US memory task (applicable to all groups except groups 9 and 10); c. Inability to adhere to the auditory instruction of selecting number 9 (applicable to groups 7 to 10); d. Having less than two correct responses to questions related to the story presented during the acquisition phase (limited to group 7); e. Taking off the headphones (>= 20% response to the question "How frequently did you remove your headphones when you heard the loud noise?"; 0% = "Never", 100% = "Always") or reducing the system volume during the experiment (relevant only to groups 9 and 10). A total of 153 persons were excluded from the statistical analysis on the basis of the above criteria, details of which are given in **Table 1**.

Table 1. Reasons for exclusion in each group.

Group	Reasons for exclusion (n)
Reference	b (2)
Fear ratings	b (2)
Expectancy ratings	b (7)
Partial reinforcement	b (8)
Trace	b (8)
CS preexposure	a (3); b (7)
Story	b (2); c (2); d (8); b & c (1); b & d (2)
Sound CS	b (1); c (11); b & c (3)
Sound US	a (1); c (6); e (35); c & e (13)
Overlap CS-US	c (1); e (24); c & e (1)
Unspecific instruction	b (3)
Specific instruction	b (2)
Total	153

Notes. a. Non-compliance with instructions in two main IMCs (applicable to all groups); b. Attaining an accuracy score of less than 80% on the US memory task (applicable to all groups except groups 9 and 10); c. Inability to adhere to the auditory instruction of selecting number 9 (applicable solely to groups 7 to 10); d. Having less than 2 correct responses to questions related to the story presented during the acquisition phase (limited to group 7); e. Taking off the headphones (>= 20% response to the question "How frequently did you remove your headphones when you heard the loud noise?"; 0% = "Never", 100% = "Always") or reducing the system volume during the experiment (relevant only to groups 9 and 10).

2.2.4 Awareness definition

We classified participants as being aware of the contingencies if they met two criteria: (i) they gave a positive response to the contingency awareness question and correctly identified the shape (or sound in Group 8) used as CS+, and (ii) they reported greater US-expectancy after the CS+ than after the CS-. Those not meeting at least one of these criteria were considered "unaware". This procedure of using two awareness tests is in line with recent recommendations on studying unaware processes (Shanks, 2017).

2.2.5 Confirmatory analyses

To explore how awareness influenced CS affective ratings (H.1), repeated measures ANOVAs were employed, followed by Holm-Sidak post-hoc tests. The post-hoc t-tests were supplemented by Bayesian t-tests which allow us to quantify support for the null hypothesis (i.e., no learning in contingency unaware participants) more directly (Dienes, 2014). The ANOVAs included within-subject factors Time (baseline, post-acquisition, and post-extinction) and CS type (CS+ vs. CS-) and between-subject factor Awareness (yes vs. no).

For comparing contingency awareness rates between groups (H.2-11), we used chi-squared tests. When dealing with three levels of the variable (Type of instruction), we initially conducted a chi-square test. If this yielded a significant result, we conducted post-hoc pairwise comparisons.

To validate the reliability of the chi-square test results, we conducted regression analyses where we controlled for potential confounding effects due to imbalanced variables between groups. Specifically, in binary logistic regression models, we used 'group' as the independent variable and 'contingency awareness' as the dependent variable, while also including significant variables identified in the homogeneity tests as covariates.

To assess the homogeneity of the groups, we compared demographic variables, personality traits, working memory, mind wandering, and US affective ratings across the groups. We employed chi-square or Fisher's exact tests for categorical variables. For normally distributed continuous variables, we used ANOVA or t-tests, while we used the Mann-Whitney U test or the Kruskal-Wallis test for non-normally distributed continuous variables. These comparisons were conducted separately for two groups corresponding to each variable. In the case where a variable (Type of instruction) had three levels, we began with an overall test. If this test yielded a significant result, we proceeded with post-hoc pairwise comparisons.

The effect of latent inhibition on average differential fear ratings (H.6.2) was assessed using a t-test. This test compared the differences in average differential fear ratings between aware participants in group 6 and group 1.

To investigate how individual differences modulate the probability to acquire awareness (H.12-16), we performed a two-step statistical analysis. In the first step, we compared demographic and psychological characteristics between aware and unaware participants. For categorical variables, we used Chi-squared or Fisher's exact tests. For normally distributed continuous variables, we employed independent samples t-tests, and for non-normally distributed continuous variables, we used Mann-Whitney U tests. In the second step, we conducted a multivariate analysis using binary logistic regression to identify independent factors that could explain variations in contingency awareness. This multivariate analysis included only variables with p-values below 0.005 in the univariate analyses.

All p-values were two-tailed, and a significance level of 0.005 was used for the tests involving all 12 groups (H.1, H.12-16). Importantly, for H.2-11 and ‘Checks of homogeneity of the groups’ the alpha level was set at 0.05. These analyses and significance levels are reported separately below to avoid confusion.

2.3 Results

2.3.1 Contingency awareness rate comparison between the groups

Participants were classified as aware of the contingencies if they correctly identified the CS+ in a two-alternative forced choice question that presented both CS+ and CS- stimuli (e.g., two shapes), and if they demonstrated greater US-expectancy following the CS+ compared to the CS-. Those who failed to meet either of these criteria were considered "unaware" (see **Table 2** for group sample sizes).

The highest contingency awareness rates were observed in groups where we employed a loud noise as the US (94.7% in the Sound US group and 96% in the Overlap CS-US group), while in the other multimodal conditioning group (auditory CS paired with visual US) the awareness rate was lower than in the reference (53.3%). The lowest rates were seen in groups that listened to a concurrent story (48%) and where fear ratings were displayed during the presentation of the conditioned stimulus (47.2%), suggesting a strong effect of distraction from the task. Interestingly, the presentation of another type of rating - expectancy - overrode the effect of diverted attention, pushing the contingency awareness rate up to 90.7%.

Adding more specificity in the instruction produced the expected effect of an increase in contingency awareness rates (unspecific: 82.7%; specific: 73.3% as compared with the reference 47.2%). Unexpectedly, however, the specific instruction effect was numerically (but not significantly) weaker as compared to the unspecific instruction.

Latent inhibition, decreasing reinforcement rate, and adding a 2-second long trace did not produce significant effects on the awareness rate as compared with the respective reference group. Similarly, the effect of introducing an overlap in time between CS and US was negligible.

Percentages of aware participants in each group are shown in **Figure 3**. The results of the chi-square tests are shown in **Table 2**.

Adding variables that differed between comparison groups as covariates (see **Appendix 6.1**: Homogeneity checks) in logistic regression models did not alter the results of the chi-square tests (detailed in the **Table 3**).

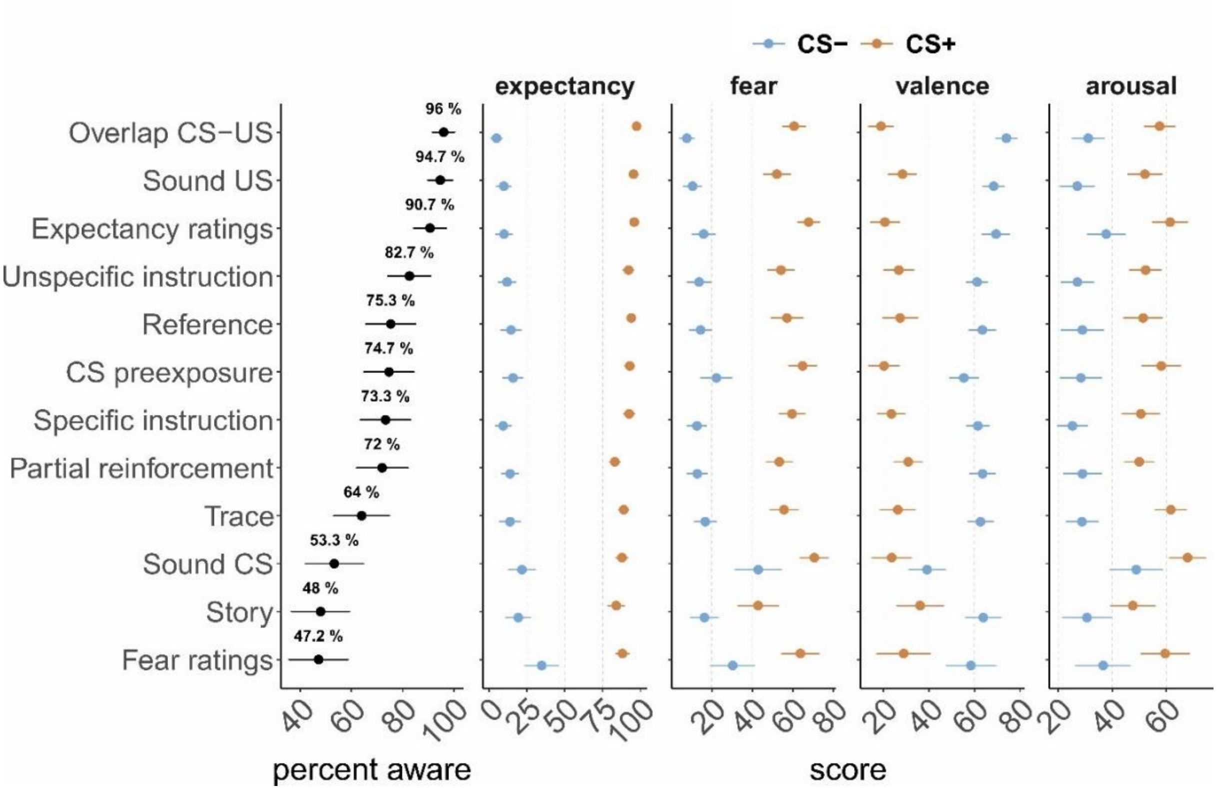


Figure 3. Percentage of aware participants in each group and post-acquisition US-expectancies and affective ratings for CS+ and CS- in aware participants in each group. Error bars are the 95% confidence intervals.

Table 2. Contingency awareness rate comparison between the groups.

Group	Aware participants	Paired group	χ^2	*p*	Cramer's *V*
Reference	75.3% (55/73)				
Fear ratings	47.2% (34/72)	Reference	12.09	**<0.001***	0.289
Expectancy ratings	90.7% (68/75)	Reference	6.19	**0.013***	0.204
Partial reinforcement	72.0% (54/75)	Reference	0.21	0.644	0.038
Trace	64.0% (48/75)	Reference	2.25	0.134	0.123
CS preexposure	74.7% (56/75)	Reference	0.01	0.924	0.008
Story	48.0% (36/75)	Reference	11.68	**<0.001***	0.281
Sound CS	53.3% (40/75)	Reference	7.80	**0.005***	0.230
Sound US	94.7% (71/75)	Reference	10.92	**<0.001***	0.272
Overlap CS-US	96.0% (72/75)	Sound US	0.15	0.699	0.032
Unspecific instruction$^{\$}$	82.7% (62/75)	Fear ratings	20.37	**<0.001***	0.372
		Specific instruction	1.90	0.168	0.113
Specific instruction$^{\$}$	73.3% (55/75)	Fear ratings	10.48	**0.001***	0.267

Notes. The group with CS and US (loud noise) overlapping in time (Overlap CS-US) used the Sound US group as its reference. Both the group given unspecific instructions and the group given specific instructions were compared with the group where concurrent fear ratings were presented (fear ratings) as their benchmark for comparison. $^{\$}$A significant difference in the rates of aware subjects was found among the groups: Fear Ratings, Unspecific instruction, and Specific Instruction ($\chi^2(2) = 22.69$, $p < 0.001$). *$p < .05$.

Table 3. Group differences in contingency awareness rates based on logistic regression.

Outcome	Factor	Covariate	Wald χ^2	OR	95% CI	*p* value
Awareness	Group (Reference vs. Fear ratings)	Smoking within two hours	10.42	0.31	0.15 - 0.62	**0.001***
Awareness	Group (Reference vs. Partial reinforcement)	age and BAS Fun Seeking	0.09	0.89	0.40 - 1.96	0.763
Awareness	Group (Reference vs. Trace)	sex, age, BAS Fun Seeking, and BAS Reward Responsiveness	1.68	0.59	0.26 -1.31	0.194

Awareness	Group (Reference vs. CS preexposure)	age and native language	0.34	1.30	0.54 - 3.15	0.560
Awareness	Group (Reference vs. Story)	age, native language and socioeconomic status	4.83	0.42	0.19 - 0.91	**0.028***
Awareness	Group (Reference vs. Sound CS)	age and native language	4.91	0.44	0.21 - 0.90	**0.027***
Awareness	Group (Reference vs. Sound US)	age, mind wandering, valence ratings of US, and arousal ratings of US	9.73	8.21	2.40 - 35.40	**0.002***
Awareness	Group (Sound US vs. Overlap CS-US)	native language, handedness, trait anxiety, state anxiety and IUS	0.56	2.00	0.33 - 14.30	0.455
Awareness	Group (Fear ratings vs. Unspecific instruction)	age, drinking coffee within two hours, native language, and US arousal ratings	18.13	8.55	3.36 - 24.66	**< 0.001***
Awareness	Group (Fear ratings vs. Specific instruction)	age, native language, BAS Drive, and US arousal ratings.	8.39	3.37	1.51 - 7.92	**0.004***

Notes. CI = confidence interval; vs. = versus; BAS = behavioral activation system; CS = conditioned stimuli; US = unconditioned stimuli; IUS = Intolerance of Uncertainty Scale. * p < .05.

2.3.2 Learning variation between groups

In our preregistration, based on the literature, we laid out the hypothesis that latent inhibition would decrease differential conditioning indicators (affective ratings). This hypothesis was not confirmed. In a set of exploratory analyses, we compared the affective ratings of CS and the differences between them (CS+ minus CS-) across all twelve groups of aware participants in the post-acquisition phase. Independent sample t-tests were carried out.

The weakest differential fear-conditioned responses were observed in Story, Sound CS, and Fear ratings, the three conditions that also had the lowest contingency awareness rates. The highest ratings were observed in the groups with CS-US overlap (having a loud noise as US) and concurrent expectancy ratings.

Despite the large difference between the awareness rates, using a loud noise as US did not produce stronger learning effects compared to the reference group as expressed in the differential affective ratings.

Using sounds as CSs significantly increased conditioned responses, as shown by higher fear ratings for both CS+ and CS-, but the difference between them was small, especially in the valence ratings.

Furthermore, incorporating online fear ratings also heightened fear responses, particularly for CS-, thus decreasing the CS+/CS- difference. Conversely, both unspecific and specific instructions decreased fear responses to CS- but did not affect the magnitude of differential conditioned responses. See **Figure 3 and Table 4** for more details.

2.3.3 Contingency awareness and learning indicators

We tested whether differential learning differed between participants aware or unaware of the contingencies in the whole sample. A highly significant interaction of Awareness x CS type x Time was observed in all learning indicators (i.e., affective ratings), including valence, arousal, and fear (all *ps* < 0.001) in the ANOVAs (refer to **Table 5**). To explore this interaction, we conducted targeted comparisons between the CS+ and CS- ratings at various phases (baseline, post-acquisition, and post-extinction), separately in the aware and unaware participant groups.

As depicted in **Figure 4,** at baseline, aware participants showed no significant difference in valence, arousal, and fear ratings between CS+ and CS- (valence: $t(893) = 1.49$, $p = 0.136$, $d = 0.05$; arousal: $t(893) = 0.02$, $p = 0.988$, $d < 0.01$; fear: $t(893) = 1.89$, $p = 0.059$, $d = 0.06$). However, in the post-acquisition phase, CS+ was rated significantly lower in valence ($t(893) = 29.35$, $p < 0.001$, $d = 0.98$), higher in arousal ($t(893) = 19.94$, $p < 0.001$, $d = 0.67$), and higher in fear ($t(893) = 32.84$, $p < 0.001$, $d = 1.10$) compared to CS-. Similar trends were observed in the post-extinction phase, with CS+ showing lower valence ($t(893) = 16.35$, $p < 0.001$, $d = 0.55$), higher arousal ($t(893) = 11.13$, $p < 0.001$, $d = 0.37$), and higher fear ratings ($t(893) = 19.73$, $p < 0.001$, $d = 0.66$) than CS-.

In contrast, the unaware participants showed no significant differences between CS+ and CS- across all three learning indicators - valence, arousal, and fear - at any of the three time points (baseline, post-acquisition, and post-extinction), with all *p*-values exceeding 0.05.

These findings are supported by Bayesian t-tests, which provided very strong evidence for differential conditioning in the aware participants and strong evidence against conditioning in the unaware participants (fear ratings after acquisition, aware: $BF_{10} = 2.41 \cdot 10^{125}$, unaware: $BF_{10} = 0.108$; arousal ratings, aware: $BF_{10} = 2.86 \cdot 10^{58}$, unaware: $BF_{10} = 0.086$; valence ratings, aware: $BF_{10} = 2.68 \cdot 10^{107}$, unaware: $BF_{10} = 0.152$).

Table 4. Differences in affective ratings of CS+ and CS- across 12 groups in aware participants after the acquisition phase.

		CSs valence ratings						CSs arousal ratings						CSs fear ratings					
		CS+		CS-		CS+- CS-		CS+		CS-		CS+ - CS-		CS+		CS-		CS+- CS-	
Group	Paired group	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*
Reference	Fear ratings	-0.28	0.779	0.88	0.385	-0.82	0.415	-1.13	0.261	-1.17	0.246	-0.06	0.956	-0.88	0.381	-2.62	**0.012***	1.12	0.268
Reference	Expectancy ratings	1.55	0.123	-1.43	0.156	1.87	0.065	-1.67	0.098	-1.62	0.107	-0.17	0.868	-1.85	0.067	-0.38	0.703	-1.43	0.155
Reference	Partial reinforcement	-0.89	0.378	-0.03	0.977	-0.53	0.597	0.26	0.799	-0.00	0.997	0.21	0.833	0.62	0.536	0.42	0.674	0.33	0.742
Reference	Trace	0.23	0.816	0.19	0.849	0.03	0.976	-1.78	0.079	0.03	0.978	-1.55	0.124	0.23	0.819	-0.59	0.558	0.60	0.552
Reference	CS preexposure	1.70	0.092	1.89	0.062	-0.17	0.865	-1.08	0.281	0.09	0.925	-0.99	0.326	-1.22	0.224	-1.68	0.096	0.01	0.991
Reference	Story	-1.71	0.091	-0.09	0.929	-1.11	0.269	0.56	0.578	-0.28	0.779	0.75	0.457	1.88	0.065	-0.44	0.660	2.10	**0.039***
Reference	Sound CS	0.70	0.484	4.79	**<0.001***	-2.74	**0.007***	-2.78	**0.007***	-3.33	**0.001***	0.46	0.647	-2.18	**0.031***	-4.81	**<0.001***	1.87	0.065
Reference	Sound US	-0.25	0.804	-1.29	0.198	0.62	0.537	-0.11	0.914	0.37	0.712	-0.37	0.713	0.80	0.426	1.07	0.285	0.17	0.863
SoundUS	Overlap CS-US	2.60	**0.010***	-1.63	0.106	2.66	**0.009***	-1.04	0.301	-0.96	0.340	-0.24	0.813	-1.59	0.114	1.03	0.306	-2.03	**0.044***
Fear ratings	Unspecific instruction	0.39	0.697	-0.49	0.626	0.63	0.528	1.05	0.299	1.57	0.122	-0.34	0.738	1.31	0.193	2.78	**0.008***	-0.86	0.392
Fear ratings	Specific instruction	1.02	0.311	-0.55	0.586	1.14	0.260	1.26	0.213	1.86	0.068*	-0.35	0.725	0.59	0.560	3.00	0.004*	-1.76	0.084
Unspecific instruction	Specific instruction	0.85	0.397	-0.10	0.922	0.63	0.528	0.30	0.767	0.42	0.674	-0.01	0.988	-0.94	0.352	0.30	0.764	-1.07	0.287

Note. CS = conditioned stimuli; US = unconditioned stimuli. * *p* < .05

Table 5. Test of Differences in Affective Ratings in the Overall Sample

Statistic		Awareness	CS type	Time	Awareness x CS type	Awareness x Time	CS type x Time	Awareness x CS type x Time
Valence	*F*	2.05	109.8	153.74	169.56	2.72	98.96	109.68
	p	0.152	**<0.001***	**<0.001***	**<0.001***	0.069	**<0.001***	**<0.001***
	η^2_p	<0.01	0.11	0.15	0.16	<0.01	0.10	0.11
	df	1.0, 893.0	1.0, 893.0	1.9, 1700.2	1.0, 893.0	1.9, 1700.2	1.9, 1675.3	1.9, 1675.3
Arousal	*F*	0.68	48.68	74.96	89.45	3.81	57.18	43.18
	p	0.408	**<0.001***	**<0.001***	**<0.001***	0.026	**<0.001***	**<0.001***
	η^2_p	<0.01	0.05	0.08	0.09	<0.01	0.06	0.05
	df	1.0, 893.0	1.0, 893.0	1.8, 1637.7	1.0, 893.0	1.8, 1637.7	1.8, 1645.6	1.8, 1645.6
Fear	*F*	3.93	163.92	250.94	228.99	3.32	133.87	135.84
	p	0.048	**<0.001***	**<0.001***	**<0.001***	0.039	**<0.001***	**<0.001***
	η^2_p	<0.01	0.16	0.22	0.20	<0.01	0.13	0.13
	df	1.0, 893.0	1.0, 893.0	1.9, 1706.7	1.0, 893.0	1.9, 1706.7	1.8, 1589.6	1.8, 1589.6

Notes. n = 895. Results of analyses of variance testing differences in affective ratings of the CSs with time (baseline, postacquisition, and postextinction) and CS type (CS+ vs. CS-) as within-subject factors and awareness (yes vs. no) as the between-subject factor. CS = conditioned stimuli. * *p* < .005.

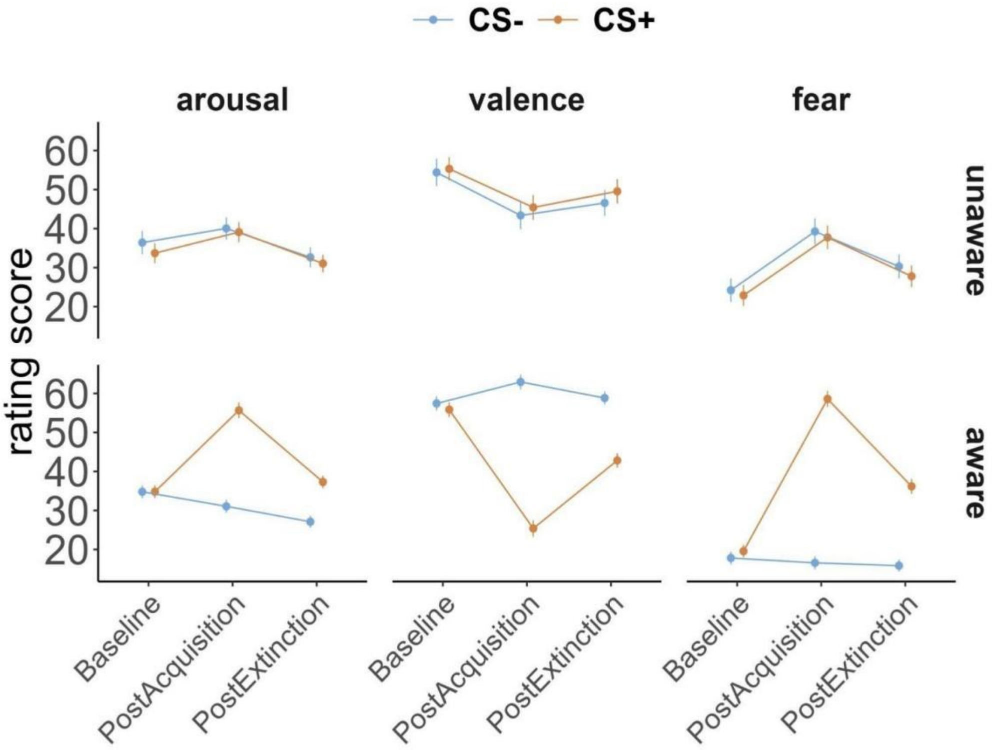


Figure 4. Mean affective ratings per time point for each CS separately among aware (*n* = 651) and unaware participants (*n* = 244). Error bars are the 95% confidence intervals.

2.3.4 Predicting contingency awareness by demographics, personality, cognitive performance and emotional response to the US

We used demographic, personality, and cognitive characteristics of the sample (see **Table 6**) to assess their contribution to the probability of becoming aware of the contingencies. In the univariate analyses, consuming alcohol within 2 hours before the experiment (12 participants in the whole sample of 895) and rating the US as more unpleasant (i.e., giving lower valence ratings) were significantly related to the probability to become aware of the contingencies (p-values below 0.005), justifying (as per preregistration) inclusion of these variables in a multiple logistic regression model. However, in this model, the effects became even smaller: 'alcohol' had a Wald χ^2 of 7.45, an odds ratio (OR) of 5.415 (95% CI: 1.610 to 18.214, $p = 0.006$), while 'valence ratings of US' had a Wald χ^2 of 4.81, an OR of 0.989 (95% CI: 0.980 to 0.999, $p = 0.028$).

The difference between contingency aware and unaware participants in their valence ratings of the US in the univariate analysis may be due to lower ratings of the loud noise as compared with the images. Indeed, the valence of the loud noise as US (groups Sound US and Overlap CS-US) was significantly lower (i.e., the noise was more unpleasant) as compared with aversive images as US (all other groups): 5.79 ± 11.21 and 17.80 ± 14.85 for the noise and the pictures, respectively, ($t(893) = 9.38$, $p < 0.001$). Similarly, the US arousal ratings of negative images scored 59.44 ± 27.08, and for loud noise as the US, the score was 73.97 ± 29.93, also indicating a significant difference ($t(893) = -5.89$, $p < 0.001$) (see **Appendix 6.6** for more detail). When we excluded the two groups with a loud noise as the US (Sound US and Overlap CS-US), the difference between contingency aware and unaware participants in their US valence ratings disappeared though all other results remained as in the original analysis (see **Appendix 6.7**).

Table 6. Comparison of characteristics between contingency unaware ($n = 244$) and aware ($n = 651$) participants.

Variables	Unaware ($n = 244$)	Aware ($n = 651$)	χ^2 / t / Z	p value	Effect size[a] ($V/d/r$)
Sex			0.37	0.545	0.020
Female	101/244 (41.4%)	255/651 (39.2%)			
Male	143/244 (58.6%)	396/651 (60.8%)			
Native Language			5.63	0.018	0.079
Other	114/244 (46.7%)	362/651 (55.6%)			
English	130/244 (53.3%)	289/651 (44.4%)			
Education			2.04	0.154	0.048
Below bachelor	76/244 (31.1%)	236/651 (36.3%)			
Bachelor or higher	168/244 (68.9%)	415/651 (63.7%)			
Alcohol[b]			-	**0.005***	0.103
No	236/244 (96.7%)	647/651 (99.4%)			
Yes	8/244 (3.3%)	4/651 (0.6%)			
Smoking			1.02	0.312	0.034
No	216/244 (88.5%)	591/651 (90.8%)			
Yes	28/244 (11.5%)	60/651 (9.2%)			
Coffee			0.49	0.486	0.023
No	157/244 (64.3%)	435/651 (66.8%)			
Yes	87/244 (35.7%)	216/651 (33.2%)			
Other drugs[b]			-	0.273	0.055
No	243/244 (99.6%)	651/651 (100.0%)			

Yes	1/244 (0.4%)	0/0 (0.0%)			
Socioeconomic			0.00	0.973	0.001
Below 6th rung	118/244 (48.4%)	314/651 (48.2%)			
6th rung or higher	126/244 (51.6%)	337/651 (51.8%)			
Age	34.35 ± 10.23	32.91 ± 10.04	1.90	0.058	0.143
Handedness	42.91 ± 10.45	43.80 ± 9.32	-1.24	0.214	-0.093
STAI trait	42.52 ± 11.40	43.02 ± 13.13	-0.53	0.596	-0.040
STAI state	37.42 ± 11.86	35.59 ± 11.84	2.06	0.040	0.155
IUS	71.40 ± 21.35	70.65 ± 20.48	0.48	0.629	0.036
BAS Drive	8.99 ± 2.55	9.27 ± 2.64	-1.42	0.155	-0.107
BAS Fun Seeking	8.75 ± 2.40	8.85 ± 2.32	-0.57	0.568	-0.043
BAS Reward Responsiveness	8.58 ± 2.93	8.30 ± 2.74	1.33	0.183	0.100
BIS	15.02 ± 4.32	14.39 ± 4.07	2.04	0.041	0.153
Working memory accuracy[c]	0.60 ± 0.13	0.62 ± 0.12	-1.99	0.047	-0.150
Mind wandering[c]	1.00 (0.00, 2.00)	1.00 (0.00, 2.00)	0.59	0.531	0.020
Valence ratings of US	16.13 (5.41, 26.25)	12.06 (1.56, 24.12)	2.88	**0.004***	0.096
Arousal ratings of US	63.55 (39.45, 81.47)	65.97 (47.29, 86.86)	-2.05	0.040	-0.069

Notes. CS = conditioned stimuli; US = unconditioned stimuli; BIS = Behavioral Inhibition System; BAS = Behavioral Activation System; STAI = State Trait Anxiety Inventory; IUS = Intolerance of Uncertainty Scale; — = no test statistic is available. [a] In the statistical analysis, Cramér's V was used to measure the effect size for the chi-square test (variables sex to socioeconomic status), Cohen's d for the t test (variables age to BIS), and r for the Mann-Whitney U test (variables working memory capacity to arousal ratings of US). Cramér's V ranges from 0 to 1, with values close to 1 suggesting a strong association and values near 0 indicating minimal or no association. Cohen's d values can extend from negative to positive infinity, with larger absolute values indicating more substantial differences between groups. For the Mann-Whitney U test, the effect size r varies from -1 to 1, with values at the extremes reflecting perfect relationships, either negative or positive, while values near 0 denote the absence of a relationship. [b] The variable was analyzed by Fisher's exact test (the same as below). Mind wandering, valence ratings of US and arousal ratings of US were reported as the median (first quartile, third quartile), and the Mann-Whitney U test was used for the analyses. The means of mind wandering were 1.44 (SD = 1.74) and 1.32 (SD = 1.58) for unaware and aware groups, respectively. The means of US valence ratings were 17.62 (SD = 14.73) and 15.10 (SD = 15.03) for unaware and aware subjects, respectively. The means of US arousal ratings were 58.54 (SD = 28.84) and 63.12 (SD = 27.72) for unaware and aware participants, respectively. [c] For the working memory accuracy and mind wandering, three participants had incomplete data, so the number of participants included in the analyses was 892. **p* < .005.

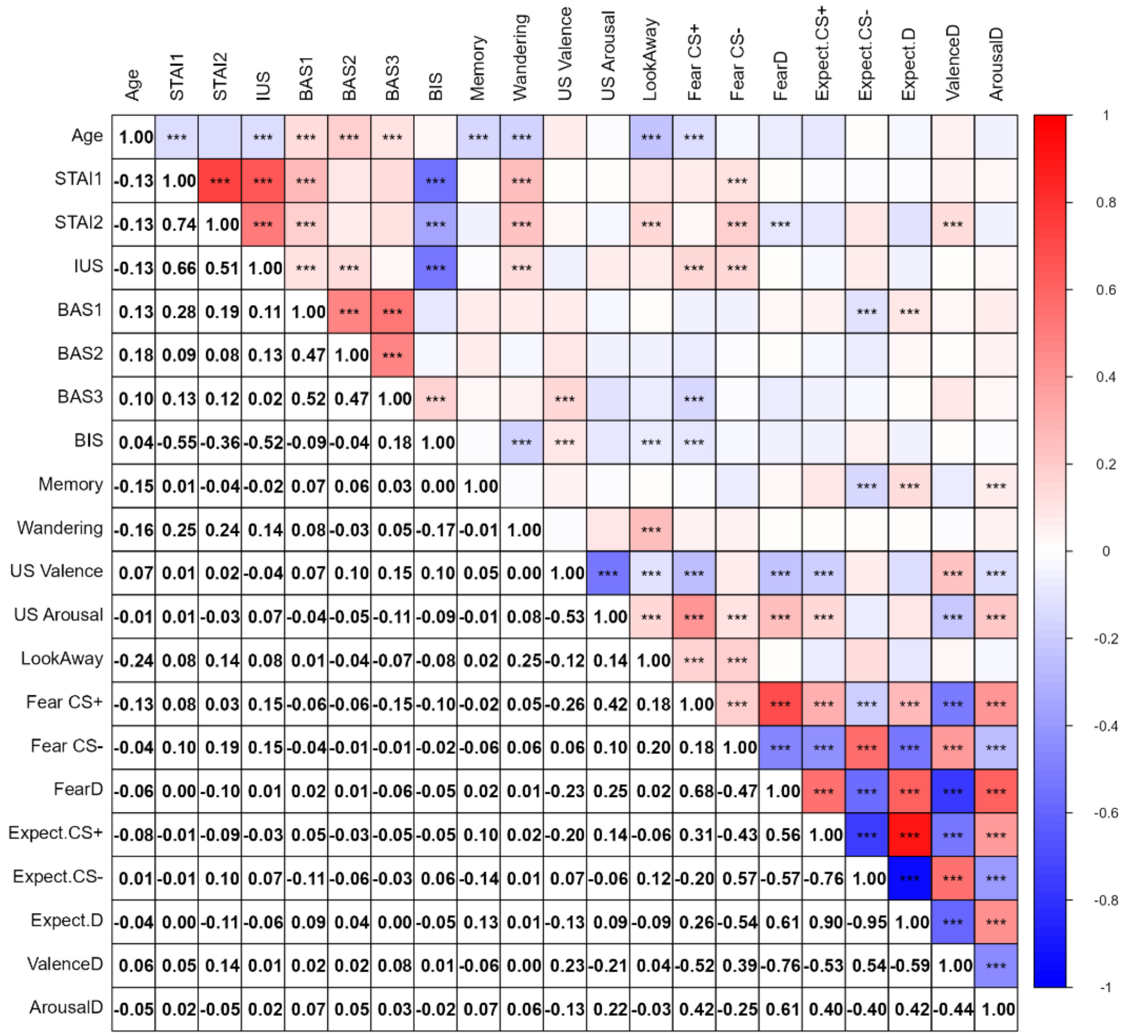
2.3.5 Correlations among the continuous variables

As a part of our exploratory analyses, we investigated the relationship between continuous differential expectancy ratings (CS+ minus CS- expectancy) post-acquisition (as an approximate continuous measure of contingency awareness) and other continuous variables (see **Figure 5**). Additionally, in our large sample, we tested the commonly observed correlations, such as the one between fear acquisition (as measured by the valence, arousal, and fear ratings) and anxiety.

Trait anxiety showed a significant positive correlation with fear ratings for CS-. Notably, the negative correlation between state anxiety and the difference in fear ratings of CS+ and CS- appears to be driven by heightened fear responses to CS- in more anxious individuals, rather than by low fear ratings of CS+ in less anxious participants. BIS score, inversely coded, was negatively and significantly correlated with fear CS+ ratings, indicating that a higher BIS level is associated with increased fear responses to CS+. IUS was positively and significantly correlated with fear ratings for both CS+ and CS-, but not with the differential score.

Avoidance behavior, as indicated by the time spent looking away from unpleasant US pictures, was significantly correlated with fearful responses towards both CS+ and CS-, but showed no correlation with differential fear ratings. This behavior also had significant associations with common measures of avoidance as a personality trait, like the BIS score. Furthermore, susceptibility to mind wandering, as a personality trait related to the ability to maintain attention to the task and resist intrinsic distracting thoughts, was positively related to the time spent looking away.

Working memory failed to show any relationship with fear and other affective ratings, yet it was significantly correlated with the differential expectancy ratings. Specifically, the better working memory was associated with a larger difference between CS+ and CS- expectancy ratings.



	Age	STAI1	STAI2	IUS	BAS1	BAS2	BAS3	BIS	Memory	Wandering	US Valence	US Arousal	LookAway	Fear CS+	Fear CS-	FearD	Expect.CS+	Expect.CS-	Expect.D	ValenceD	ArousalD
Age	1.00	***		***	***	***	***		***	***			***	***							
STAI1	-0.13	1.00	***	***	***			***		***					***						
STAI2	-0.13	0.74	1.00	***	***			***		***			***		***	***				***	
IUS	-0.13	0.66	0.51	1.00	***	***		***		***				***	***						
BAS1	0.13	0.28	0.19	0.11	1.00	***	***											***	***		
BAS2	0.18	0.09	0.08	0.13	0.47	1.00	***														
BAS3	0.10	0.13	0.12	0.02	0.52	0.47	1.00	***			***			***							
BIS	0.04	-0.55	-0.36	-0.52	-0.09	-0.04	0.18	1.00		***	***		***	***							
Memory	-0.15	0.01	-0.04	-0.02	0.07	0.06	0.03	0.00	1.00									***	***		***
Wandering	-0.16	0.25	0.24	0.14	0.08	-0.03	0.05	-0.17	-0.01	1.00			***								
US Valence	0.07	0.01	0.02	-0.04	0.07	0.10	0.15	0.10	0.05	0.00	1.00	***	***	***		***	***			***	***
US Arousal	-0.01	0.01	-0.03	0.07	-0.04	-0.05	-0.11	-0.09	-0.01	0.08	-0.53	1.00	***	***	***	***	***			***	***
LookAway	-0.24	0.08	0.14	0.08	0.01	-0.04	-0.07	-0.08	0.02	0.25	-0.12	0.14	1.00	***	***						
Fear CS+	-0.13	0.08	0.03	0.15	-0.06	-0.06	-0.15	-0.10	-0.02	0.05	-0.26	0.42	0.18	1.00	***	***	***	***	***	***	***
Fear CS-	-0.04	0.10	0.19	0.15	-0.04	-0.01	-0.01	-0.02	-0.06	0.06	0.06	0.10	0.20	0.18	1.00	***	***	***	***	***	***
FearD	-0.06	0.00	-0.10	0.01	0.02	0.01	-0.06	-0.05	0.02	0.01	-0.23	0.25	0.02	0.68	-0.47	1.00	***	***	***	***	***
Expect.CS+	-0.08	-0.01	-0.09	-0.03	0.05	-0.03	-0.05	-0.05	0.10	0.02	-0.20	0.14	-0.06	0.31	-0.43	0.56	1.00	***	***	***	***
Expect.CS-	0.01	-0.01	0.10	0.07	-0.11	-0.06	-0.03	0.06	-0.14	0.01	0.07	-0.06	0.12	-0.20	0.57	-0.57	-0.76	1.00	***	***	***
Expect.D	-0.04	0.00	-0.11	-0.06	0.09	0.04	0.00	-0.05	0.13	0.01	-0.13	0.09	-0.09	0.26	-0.54	0.61	0.90	-0.95	1.00	***	***
ValenceD	0.06	0.05	0.14	0.01	0.02	0.02	0.08	0.01	-0.06	0.00	0.23	-0.21	0.04	-0.52	0.39	-0.76	-0.53	0.54	-0.59	1.00	***
ArousalD	-0.05	0.02	-0.05	0.02	0.07	0.05	0.03	-0.02	0.07	0.06	-0.13	0.22	-0.03	0.42	-0.25	0.61	0.40	-0.40	0.42	-0.44	1.00

Figure 5. Correlation matrix of continuous variables. Notes. The labels defined in the picture are as follows: STAI1 denotes trait anxiety, and STAI2 represents state anxiety. BAS1, BAS2, and BAS3 correspond to BAS Drive, BAS Fun Seeking, and BAS Reward Responsiveness, respectively. 'Memory' refers to the accuracy of working memory, 'Wandering' to the incidence of mind wandering during the working memory task, and 'LookAway' to the frequency of looking away from the screen during the presentation of unpleasant pictures. Fear measurements include FearCS+ and FearCS- for the fear ratings of CS+ and CS-, respectively, with FearD capturing the difference between these ratings. For US-expectancies, ExpectCS+ and ExpectCS- pertain to CS+ and CS-, respectively, with ExpectancyD illustrating the differences between these expectancies. ValenceD and ArousalD are used to describe the differences in valence and arousal ratings between CS+ and CS-, respectively. Assessments for FearCS+, FearCS-, FearD, ExpectCS+, ExpectCS-, ExpectancyD, ValenceD, and ArousalD all took place after the acquisition phase. Participant counts were as follows: 892 for working memory and mind wandering, 745 for looking away, and 895 for other correlations. All correlations are Spearman's rank correlation coefficients. ***$p < 0.005$.

2.4 Discussion

In this study, we investigated the impact of contingency awareness on fear acquisition learning across a large online sample of adults. The effect was clear and unequivocal: successful fear acquisition was observed exclusively among participants who reported explicit contingency awareness, as evidenced by significant changes in affective responses to the CS, suggesting that fear acquisition learning does not occur without contingency awareness. Furthermore, we identified several experimental factors that affected the acquisition of contingency awareness: the presence of online fear ratings, the presence of online US-expectancy ratings, engagement in concurrent tasks, the sensory modality of the CS, the sensory modality of the US, and the instructions provided. Lastly, we found no evidence that personality traits, cognitive abilities, or demographic factors moderate the process of contingency awareness acquisition.

Conditioned affective responses, measured by self-report, were observed only in participants who were aware of the contingencies. This finding aligns with the outcomes of some prior studies (Baeuchl et al., 2019; Glenn, Klein, et al., 2012; Klucken et al., 2009; Tabbert et al., 2011), but is in contrast to others that suggested that fear acquisition learning may not necessitate awareness of contingency (Raio et al., 2012; Schultz & Helmstetter, 2010; Steinberg et al., 2012). We believe that many previous studies that demonstrated unaware learning did not carefully filter out aware participants from the unaware group or condition. In a seminal review, Lovibond and Shanks pointed out that numerous studies on conditioning have employed inadequate measures to assess contingency awareness, thus questioning the validity of findings supporting conditioning without contingency awareness (Lovibond & Shanks, 2002). Likewise, a meta-analysis revealed that a significant proportion of studies on contingency unaware fear conditioning suffered from methodological limitations; particularly, studies employed poor awareness measures and provided the researchers with too many degrees of freedom (Mertens & Engelhard, 2020). Moreover, this analysis indicated that higher-quality studies were less likely to affirm the possibility of conditioning in the absence of contingency awareness. Our study adds more high-quality evidence against the possibility of fear acquisition learning in a fear conditioning paradigm without contingency awareness. In this context, the literature on evaluative conditioning, mentioned above, may provide additional insights. In evaluative conditioning paradigms, pairing a CS with another emotional outcome (co-occurrence) prompts changes in the affective evaluation of the CS, and it is increasingly accepted that this process requires contingency awareness (Moran et al., 2023), paralleling the findings reported here. Additional similarities between our approach and traditional evaluative conditioning paradigms may be explored systematically in future research.

We varied several design parameters between 12 groups of participants to understand which parameters have the potency to modulate contingency awareness. Summarising the evidence, direction of attention during conditioning was the most crucial factor affecting awareness. Introducing online fear ratings or instructing participants to listen to a concurrently presented story during the acquisition phase with the goal of diverting their attention away from the CS-US relationship hindered the acquisition of awareness. In contrast, the use of instructions containing contingency information or the implementation of online US-expectancy ratings during acquisition appeared to direct their attention more strongly toward the CS-US relationship. While US-expectancy ratings might act as a distraction - similar to fear ratings - they may instead counteract that effect by explicitly orienting participants’ attention to the CS-US contingency, thus potentially enhancing awareness. By comparison, fear ratings could shift attention primarily to the CS and participants’ subjective emotional state, rather than to the external associative link between CS and US. This shift in focus from external contingencies to internal emotional states may help explain why fear ratings seemed more disruptive to awareness formation. Interestingly, stronger facilitating learning effects of unspecific instructions (e.g., "one of the shapes will always be followed by an unpleasant picture. Your task is to find out which one") as compared with maximally specific instructions (e.g., "the circle will always be followed by an unpleasant picture and the diamond will not") has been shown previously (Mertens et al., 2021). This observation is consistent with the stated role of attention in steering the cognitive processing system toward a specific subset of available information (McDowd, 2007). Unspecific instruction directs participants’ attention to the stimuli, whereas complete, exhaustive a-priori information may diminish the engagement of active attention mechanisms. Some previous studies directly manipulated contingency awareness by altering participants' attention. For example, Kattner (2012) diverted participants' attention away from either the stimulus itself or from the CS-US relationship through two distinct distraction tasks (full-distraction task or contingency-distraction task), demonstrating that both can effectively decrease the acquisition of contingency awareness (Kattner, 2012). Other studies shifted participants' attention through instructions, such as partially or fully informing them of the CS-US relationship before the fear acquisition phase, which has also been proven to be an effective method for altering contingency awareness rate (Dawson & Biferno, 1973; Mertens et al., 2021; Tabbert et al., 2011). Altogether, available evidence underscores the relationship between the direction of attention and the acquisition of contingency awareness.

The sensory modalities of the CS and US were identified as significant experimental variables that substantially influence the acquisition of awareness. Auditory CSs, compared to visual CS, appeared to decrease the likelihood of acquiring contingency awareness. This reduction may stem from the increased difficulty in perceiving and distinguishing auditory stimuli as opposed to visual stimuli, which in turn impedes the acquisition of awareness. Alternatively, the auditory stimuli may have distracted participants from the screen. However, using loud noise as the US, instead of unpleasant images, enhanced the acquisition of awareness. The loud noise provoked significantly more negative emotional responses than the negative images and, being hard to ignore without breaching the task requirements, likely compelled participants to be attentive to the noise and to recognize its predictive signal (CS+), thus facilitating the acquisition of contingency awareness.

Among other factors, pre-exposure to CSs and US can influence subsequent conditioning processes. Research involving both human and animal subjects has demonstrated that such pre-exposure leads to latent inhibition, which is marked by reduced conditioned responses (Jordan et al., 2015; Meulders et al., 2012; Vaitl & Lipp, 1997). Contrary to the expectation, our study did not observe a reduction in conditioning after the habituation phase. Although this result suggests that the previously documented latent inhibition effect (primarily shown in animal studies) may not generalize straightforwardly to our human paradigm, we must exercise caution in interpreting the lack of significant differences. Given that our study was powered to detect medium-sized effects, it is possible that smaller yet meaningful effects of latent inhibition on fear acquisition learning remained undetected.

When examining the most effective differential conditioning, we observed that the paradigm with overlapping CS and loud noise US demonstrated the strongest conditioning effects. This should be viewed in the context of many participants being excluded due to a failure to follow instructions, such as altering the US stimulus parameters to make them less aversive or avoiding it entirely - behaviours enabled in the online setting. With aversive images as US, the factors contributing to awareness were unspecific instructions (which encourage curiosity and exploration in the CS-US association) and the use of expectancy ratings (which directs participants' attention to the fact that the US appears only after a particular CS). A combination of giving instruction, using a strongly aversive US, and the absence of distractors is thus expected to result in maximizing contingency awareness along with increasing learning effects, further emphasizing their strong relationship.

If we refrain from interpreting the result that twelve participants who had consumed alcohol before the experiment were less likely to become aware of the contingencies, and if we concentrate on the remaining 98.7% of participants, we found no significant impact of the inter-individual variability in demographics, personality traits, and cognitive abilities on the acquisition of contingency awareness. First of all, taking into account our sample size, this finding differs from similar results in the literature because it cannot be attributed to false negatives. Furthermore, the variables used to characterize participants were not invalid in general, as they correlated with fear acquisition learning in a predictable way: Anxiety, behavioural inhibition, and avoidance behaviour were significantly related to fear ratings in conditioning experiments. However, neither these nor other variables increased or decreased the probability of becoming aware of the contingencies. One may speculate that the ability to explicitly formulate the acquired knowledge of a new association is a specific feature possibly related to verbal abilities. Although working memory generally correlates with verbal intelligence, the WM test used in the present study may be too unspecific in this relation. From this point of view, not only performance in specific verbal intelligence tasks would be of interest but also hypothesis-driven fMRI experiments looking for links between contingency awareness and the activity in the left frontal and temporal areas.

A key limitation of our study is that we used a single modality of measurement - self-reports - to infer fear learning. Although fear conditioning research often employs multiple indices of learning, including physiological measures such as skin conductance responses (SCR), the startle reflex, or pupil size, implementing these online remains challenging. Nonetheless, prior work suggests that self-report and physiological measures show substantial convergence in standard fear conditioning paradigms, demonstrating comparable sensitivity to experimental manipulations such as verbal instructions or difficult-to-discriminate CSs (Mertens & De Houwer, 2016; Singh et al., 2013). Hence, we believe our primary findings are unlikely to be entirely explained by demand characteristics. Future research could explore the integration of physiological markers, such as pupil dilation or heart rate, into online experiments to enhance the validity of the findings.

Furthermore, although online data collection enables recruiting large and diverse samples, it inevitably reduces direct experimental oversight compared to laboratory settings. Online participants, in contrast, work in uncontrolled, variable environments, making it difficult to confirm whether they consistently follow instructions or remain attentive throughout the session. Nonetheless, many in-lab studies face similar issues: participants may daydream, become distracted, or fail to engage fully with the experimental task, even if they remain seated before an experimenter. In both online and in-person designs, researchers must rely on indirect indicators of attention, such as reaction times, performance on filler tasks, or self-reports of engagement. While such checks do not fully replicate the experimenter's ability to observe body language or eye gaze in real time, they help identify participants who are clearly disengaged, but these observations are rarely done in lab experiments either as the experiments have become much more automated and experimenter-independent. In future work, researchers could employ additional methods, such as requiring webcam-based gaze tracking or introducing interactive trial-by-trial feedback, to further ensure compliance.

Despite these limitations, our findings - together with those of other online studies (Bridges et al., 2020; Sauter et al., 2022) - suggest that reliable data can be collected remotely if appropriate safeguards are put in place. To mitigate these concerns, we included sound calibration and attention checks tailored to our specific tasks, particularly in the sound-based conditions. Among 1048 participants we tested, only 4 did not pass IMC. We also assured that the loud noise used in our online study was similar to the one we commonly use in lab-based fear conditioning experiments. Our data indicate that participants rated both the loud noise and aversive images as highly unpleasant, confirming that the US was sufficiently aversive in the online setting. When paired with careful technical, attention, and data-quality checks, online testing provides a powerful, efficient way to conduct large-scale experimental research - an important consideration given persistent concerns about underpowered and non-diverse samples in psychology.

2.5 Conclusion

The present study investigated the relationship between contingency awareness and fear acquisition learning in a fear conditioning paradigm, providing decisive support for the notion that the manifestation of conditioned fear in affective ratings of the CS necessarily requires awareness of the relationship between the CS and US. This finding reinforces the idea that fear acquisition learning relies on a single propositional learning mechanism (Lovibond & Shanks, 2002), thereby deepening our comprehension of how people learn from fear-related experiences. Highlighting the significant impact of experimental settings on the development of awareness, the application of specific changes to the experimental paradigm alters the contingency awareness rate, ranging from 47% to 96%, with the strongest effects being exerted by manipulations of attentional focus. This underscores the importance of carefully choosing experimental parameters to align with a given study's objectives. For example, in future studies, to reduce the likelihood of participants becoming aware of the contingencies, incorporating distraction tasks is an effective strategy. Conversely, to promote contingency awareness, offering instructions that encourage curiosity or employing a stronger US can be effective. Lastly, our findings suggest that while individual differences should not be overlooked as they play a significant role in shaping fear conditioning outcomes (Lonsdorf & Merz, 2017), the role of individual differences in predicting contingency awareness is negligible.

3 Study 2: Differential Effects of Unconditioned Stimulus Modality on Aversive Conditioning: Self-Report and Neurophysiological Evidence

3.1 Introduction

Within classical conditioning, the unconditioned stimulus (US) is central to shaping both the rate of acquisition and the strength of conditioned responses (CRs), as proposed in the Rescorla-Wagner model (Rescorla & Wagner, 1972). However, systematic comparisons of different US modalities in human aversive conditioning remain scarce, limiting our understanding of how US characteristics shape learning outcomes.

Among available US modalities, electric shocks remain the most commonly employed US in human aversive conditioning studies (Fullana et al., 2016; Mertens & Engelhard, 2020). As a nociceptive stimulus, they consistently activate the defensive system and promote robust fear acquisition when paired with a neutral cue (Wiech & Tracey, 2013). Their popularity is further reinforced by their high experimental controllability, including precise timing and individually calibrated intensity for each participant.

Aversive auditory stimuli such as loud noises (Abramson et al., 2024; Reinhard et al., 2024; Yin et al., 2020), human screams (Björkstrand et al., 2022; Lambert et al., 2021; Poplin et al., 2024), and unpleasant metallic scraping sounds (Neumann & Waters, 2006) have also been shown to elicit conditioned fear responses. Comparisons across modalities have shown that conditioned stimuli (CSs) paired with electric shock often elicit stronger fear-potentiated startle and skin conductance responses (SCRs) than those paired with a fearful scream, despite participants giving comparable subjective ratings of unpleasantness to the CSs (Ney, Nichols, et al., 2023). In contrast, a study reported that a loud, aversive scraping sound elicited CRs in SCRs and heart rate (HR) that were comparable to those evoked by electric shocks and loud tones (Neumann & Waters, 2006). Furthermore, when an extended number of acquisition trials were used, loud white noise elicited stronger CRs than electric shocks, as reflected in SCRs, HR changes, and subjective affective ratings (Sperl et al., 2016). Taken together, the findings imply that while electric shocks generally elicit stronger CRs than aversive sounds, certain auditory stimuli can produce comparable or even stronger effects under specific conditions.

In addition to auditory and nociceptive USs, visual US such as aversive images (Raes et al., 2009; Schweckendiek et al., 2011) and negative film clips (Hauck et al., 2022; Lam et al., 2024) constitute another commonly employed modality in aversive conditioning, although they are generally regarded as less intense. A recent meta-analysis found that aversive film clips can elicit CRs similar to those produced by electric shocks, as reflected in fear-potentiated startle responses, US expectancy ratings, and CS unpleasantness ratings. However, these visual stimuli tend to evoke weaker SCRs (Ney, Schenker, et al., 2022). Importantly, the meta-analysis did not include direct within-study comparisons between aversive film clips and electric shocks, highlighting the need for further empirical investigation. Similarly, studies directly comparing the differential effects of aversive images and electric shocks as US on CRs remain limited, indicating a significant gap in the current literature.

Tactile stimulation, particularly airpuffs, is commonly employed in classical eyeblink conditioning paradigms (Allen et al., 2018; Weidemann et al., 2016). Beyond their conventional use in eyeblink conditioning, airpuffs have been used as US in human aversive learning paradigms, given their cross-species aversive properties (Lovett-Barron et al., 2014; Pine et al., 2001). For instance, one study paired a throat-directed airpuff with a colored visual cue (CS+), and found increased BOLD signal changes in the right amygdala for the CS+ as opposed to the CS- (Pine et al., 2001). In animal research, high-intensity corneal airpuffs have been shown to support HR conditioning, producing effects comparable to those elicited by medium- or high-intensity periorbital shocks (McEchron et al., 1992). Despite these findings, the effectiveness of eye-directed airpuffs as aversive US in humans remains insufficiently understood, particularly when compared to electric shocks or aversive images in terms of multiple CRs.

As outlined above, different US modalities can elicit different CRs, which may be partly attributable to the distinct emotions they induce. Electric shocks, due to their painful nature, predominantly evoke fear-related reactions. In contrast, negative images tend to induce a broader spectrum of emotions, with fear and disgust being the most common (Marchewka et al., 2014). Airpuffs are generally perceived as mildly aversive and typically produce a negative affective state characterized by displeasure and startle-like arousal (Kauvar et al., 2025), possibly involving low-level disgust rather than intense fear. Importantly, these emotional responses are not merely subjective but correspond to distinct autonomic nervous system profiles. Fear is primarily associated with sympathetic activation, such as increased SCRs, which facilitates defensive responses like fight or flight. Disgust, in comparison, is more closely linked to parasympathetic activity and is often accompanied by HR deceleration and overall physiological inhibition (Klucken et al., 2012). These patterns suggest that different US modalities may recruit distinct autonomic systems, contributing to variation in conditioned physiological responses.

In line with these autonomic distinctions, neuroimaging evidence indicates that fear and disgust are supported by partially distinct neural systems. Whereas disgust tends to recruit the insula, visual regions, and frontal cortex, fear is more reliably associated with heightened activity in the amygdala, medial prefrontal cortex and hippocampus (Kirby & Robinson, 2017; Lindquist et al., 2012; Murphy et al., 2003; Pujol et al., 2018; Vytal & Hamann, 2010). These findings imply that different US modalities may engage distinct neural pathways during aversive learning, resulting in variability in conditioned neural responses. However, despite this growing evidence, relatively few studies have directly compared how different US modalities modulate neural activity during aversive conditioning.

Electroencephalography (EEG) offers a valuable method for capturing such differences owing to its fine temporal resolution and sensitivity to anticipatory and evaluative processes. Several EEG components are particularly relevant in this context. The stimulus-preceding negativity (SPN) reflects anticipatory attention and motivational significance prior to an expected outcome (Böcker et al., 2001; Ferreira de Sá et al., 2019). The late positive potential (LPP) reflects sustained evaluative processing of emotionally salient stimuli and is generally enhanced by emotionally charged content (Bacigalupo & Luck, 2018; Panitz et al., 2015; Sperl et al., 2021). Additionally, reductions in alpha-band power reflect heightened cortical excitability and increased attentional engagement, and have frequently been observed during aversive learning paradigms (Bacigalupo & Luck, 2022; Farkas et al., 2024). Together, these EEG markers provide a window into how US modality shapes anticipatory and evaluative processes during aversive learning.

Building on the above findings, the present study aimed to directly compare three US modalities: electric shocks, aversive images, and airpuffs, within a unified aversive conditioning paradigm which was directly built using the design of Study 1. We hypothesized that these US will elicit distinct patterns of CRs across subjective, autonomic, and neural systems. Specifically, we expected electric shocks to produce the strongest conditioned affective responses, followed by aversive images, while airpuffs are expected to elicit the weakest responses. In terms of physiological outcomes, electric shocks are predicted to evoke the highest conditioned SCR, reflecting strong sympathetic activation. Airpuffs, in contrast, are expected to result in greater conditioned HR deceleration, consistent with parasympathetic engagement. Aversive images may produce more variable conditioned physiological effects due to their diverse emotional content. Finally, at the neural level, EEG measures such as SPN, LPP, and alpha power suppression are expected to show modality-specific differences, in line with variations in conditioned anticipation and emotional processing. Understanding how different US modalities influence CRs across affective, physiological, and neural systems can deepen our understanding of aversive learning mechanisms and inform the selection of appropriate US in experimental design.

A secondary objective of this study was to examine the role of contingency awareness in fear conditioning. In Study 1, a large-scale online study, we found that conditioned affective responses emerged only among participants who demonstrated awareness of the CS-US contingency. The current study is designed to replicate and extend this finding under controlled laboratory conditions. Specifically, within the image US group, we examine not only the relationship between contingency awareness and conditioned affective responses, but also its association with conditioned physiological responses, including HR and SCRs. We hypothesize that significant CRs, as indicated by affective ratings, SCRs, and HR changes, will be observed exclusively in participants who show awareness of the CS-US relationship. The findings from this analysis are expected to illuminate the role of contingency awareness in fear learning.

3.2 Methods

3.2.1 Participants

Participants were recruited via the University of Tübingen email system. Participants were eligible if they were between 18 and 59 years of age and satisfied further criteria: absence of neurological or psychiatric history, not being pregnant or breastfeeding, adequate English proficiency, and normal or corrected-to-normal vision. The experimental session lasted approximately 2 to 2.5 hours. Participants were compensated with either €25 or course credit.

In total, 116 participants fulfilled the eligibility requirements and participated in the experiment. Due to issues with the experimental procedure, data from one participant were excluded, resulting in a final sample of 115 participants (age: *M* = 25.23, *SD* = 4.60, range = 19-41 years; 71.30% female). Participants were randomly allocated to one of three experimental groups, each defined by a distinct sensory modality of the unconditioned stimulus (US): aversive images, electric shocks, or airpuffs. Group 1 (*n* = 52; age: *M* = 25.98, *SD* = 5.28; 69.23% female) received aversive images as the US. Group 2 (*n* = 28; age: *M* = 24.00, *SD* = 3.43; 78.57 % female) received electric shocks as the US. Group 3 (*n* = 35; age: *M* = 25.09, *SD* = 4.20; 68.57 % female) received airpuffs as the US. The three groups did not differ significantly with respect to age or sex distribution (age: $F(2, 112) = 1.73$, $p = 0.181$; sex: $\chi^2(2) = 0.96$, $p = 0.619$), indicating that the groups were comparable.

In addition to sex and age, we collected and compared a range of other sociodemographic characteristics and personality traits (see **Questionnaires** section below) across the three groups. The vast majority of these variables showed no significant differences between groups, further supporting the overall homogeneity of the sample and the comparability of the experimental groups (see **Appendix 6.8**).

Before the experiment began, participants received a general overview of the procedure. However, they were not informed about the relationship between the CS and the US. All participants provided informed consent and were also informed of their right to withdraw at any time without consequences. After the experiment, participants participated in a debriefing session. All data were anonymized, and the study was approved by the Ethics Committee of the University of Tübingen.

3.2.2 Procedures

Participants first completed a series of questionnaires assessing demographic variables and personality traits. They then took part in a conditioning task during which data were collected across three domains: subjective experience, physiological responses, and neurophysiological activity.

Subjective measures included self-reported US expectancy ratings and CS affective evaluations. Physiological measures comprised skin conductance responses (SCRs) and heart rate (HR), derived from electrocardiogram (ECG) recordings. Neurophysiological activity was recorded using electroencephalogram (EEG), capturing central nervous system responses during the task. For participants in the airpuff group, two additional physiological measures were also recorded throughout the task: pupil diameter and eye blink rate. After the conditioning task, all participants completed a structured debriefing session.

Throughout the experiment, participants sat comfortably in an armchair with armrests. The experiment was conducted in a soundproof, electrically shielded cabin equipped with a 22-inch LG Flatron L227WTP-PF monitor positioned about 1 meter from participants' eyes. Ambient lighting was maintained at 5 lux to ensure a consistent visual environment across participants.

Questionnaires. Demographic data were obtained from participants, covering sex, age, educational background, native language, substance use within two hours prior to the experiment (e.g., smoking, alcohol, coffee), socioeconomic status, and handedness. Personality traits were assessed using three standardized questionnaires: the Intolerance of Uncertainty Scale (IUS), which consists of 27 items (Freeston et al., 1994), the State-Trait Anxiety Inventory (STAI), which includes 40 items (Spielberger et al., 1983) and the Behavioral Inhibition System/Behavioral Activation System Scales (BIS/BAS), which include 20 items (Carver & White, 1994).

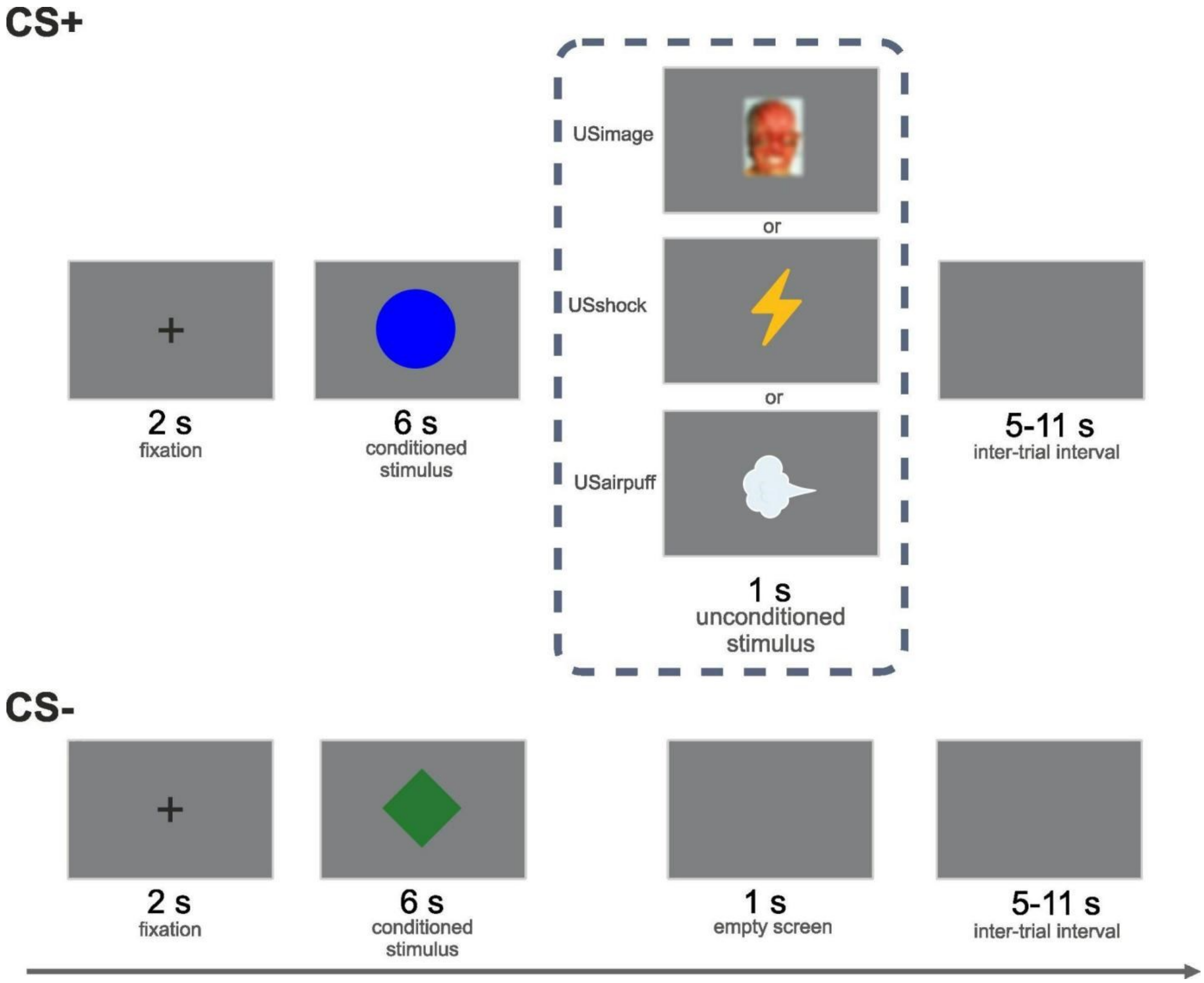


Figure 6. Schematic of the conditioning paradigm. Participants were assigned to one of three groups, each receiving a different unconditioned stimulus (US): aversive images, electric shocks, or airpuffs. All other task parameters were identical across groups. The interval between the onset of one trial and the onset of the next ranged from 14 to 20 seconds, comprising a 2-second fixation cross, a 6-second conditioned stimulus (CS), a 1-second US period, and a variable blank screen lasting 5-11 seconds. The 1-second US period involved either an image presentation (image group), a 20 ms electric shock followed by a 980 ms blank screen (shock group), or a 50 ms airpuff followed by a 950 ms blank screen (airpuff group).

Conditioning task. Participants in three groups experienced different modalities of the US: aversive images, electric shocks, or airpuffs. Apart from this difference, all other parameters of the conditioning task were identical across groups. The conditioning paradigm and experimental design are illustrated in **Figure 6**.

Group 1: Aversive images as US

In Group 1, participants were presented with aversive images as the US, while two geometric shapes (a circle and a diamond) served as the CS. One shape was randomly assigned as CS+ (always paired with the US), and the other as CS- (never paired with the US). The colors of the shapes (green and blue) were counterbalanced across participants. The CSs were presented centrally on a grey background and subtended approximately 5.1° × 5.1° of visual angle (8.88 × 8.88 cm at a viewing distance of 100 cm). The US consisted of ten unpleasant images selected from the Nencki Affective Picture System (NAPS) (Marchewka et al., 2014), based on normative ratings of fearfulness and disgust. The task design closely replicated the procedure used in the reference group of Study 1 (Li et al., 2025), with one key modification: the number of blocks was increased from one to five to meet trial number requirements for EEG analysis.

The task was subdivided into three phases: baseline, acquisition, and extinction. During the baseline phase, participants first rated a neutral white square to familiarize themselves with the rating scale. They then provided affective ratings for the circular- and diamond-shaped images. Affective ratings assessed three dimensions: valence ("How pleasant is the feeling that this picture evokes in you?"), arousal ("How intense or arousing is the feeling that this picture evokes in you?"), and fear ("How anxious does this shape currently make you feel?").

In the acquisition phase, participants underwent a total of 110 trials, subdivided into five blocks, each containing 22 trials (11 trials per CS). Each trial started with a 2-second fixation cross. Next, the CS was presented for 6 seconds. A CS+ was always followed by a 1-second presentation of a US (100% reinforcement rate), whereas a CS- was followed by a 1-second blank screen. Because each block contained 11 CS+ trials but only ten unique aversive images, one image was presented twice in each block, while the remaining nine images were each presented once. Each trial was followed by a 5 - 11 second inter-trial interval filled with a blank screen. No more than two consecutive trials involved the same CS. All stimuli were presented on a grey background.

After each block, participants answered questions about their awareness of the CS-US relationship. They were asked, "Did you notice that an unpleasant picture was shown after seeing a specific colored shape?" If they answered "Yes," they were shown both CS shapes side by side and asked, "Which shape was followed by an unpleasant picture?" They then provided US expectancy ratings, i.e., rated the likelihood of an unpleasant picture following each CS on a scale from 0 ("Certainly not") to 100 ("Certainly"), with 50 ("Uncertain"). Finally, they repeated the affective ratings (valence, arousal, and fear) for the two CS, presented in random order.

The extinction phase consisted of 14 trials (7 per CS), during which no US were presented. Participants were not informed of this change. Following the extinction phase, they completed US expectancy ratings and repeated the affective ratings. Finally, participants rated each of the ten aversive images (US) on valence and arousal, using the following questions: "How pleasant is the feeling that this picture evokes in you?" and "How intense or arousing is the feeling that this picture evokes in you?". Notably, the mean rating across the ten images was calculated separately for valence and arousal to represent the overall affective evaluation.

Group 2: Electric shocks as US

In Group 2, the procedure followed the same steps as Group 1, except that electric shocks were used as the US. Before the start of the experiment, shock intensity was calibrated to each participant’s pain threshold. A single 20 ms electric shock, generated by a Digitimer DS7A electrical stimulator, was delivered via an electrical stimulation band attached to the participant’s left wrist, positioned on top of the arm.

The calibration process began with the sensory threshold identification. Starting at 0 mA, the stimulation current was increased in 1 mA increments until the participant reported feeling the pulse. This intensity was recorded as the sensory threshold. Next, the pain threshold was identified by increasing the current in 4 mA increments from the sensory threshold until the participant reported feeling pain. To ensure reliability, both thresholds were measured twice, and the final shock intensity was calculated as the average of the two pain threshold values. The average shock intensity used in this experiment was 13.74 mA (*SD* = 7.31), ranging from 5.50 to 35.50 mA. Participants also rated the subjective intensity of the shock, yielding an average score of 4.48 (*SD* = 1.81) on a scale from 0 to 10, where 0 indicates "no pain," 5 indicates "moderate pain," and 10 indicates "worst pain."

After the extinction phase, participants rated the electric shock (US) on valence and arousal after a single presentation, using the following questions: “How pleasant is the feeling that this electric pulse evokes in you?” and “How intense or arousing is the feeling that this electric pulse evokes in you?”.

Group 3: Airpuffs as US

In Group 3, the procedure was identical to that of Group 1 and Group 2, except that airpuffs served as the US. Specifically, the US was a 50 ms airpuff delivered through tubing with an inner diameter of 4 mm and narrowing to a tube with a 2 mm inner diameter at the point of air release. The valve pressure was set to 1 bar and controlled by an Arduino, which is an open-source microcontroller platform commonly used for hardware automation and prototyping. The actual output pressure at the eye was 0.5 bar (approximately 7.25 psi). The tube, attached to an adjustable head-mounted headset, was positioned 1 to 2 cm from the right corner of the participant’s left eye. Throughout the experiment, ocular responses were continuously monitored and recorded via an eye tracker.

For calibration, the airpuff was initially delivered to the participant's closed eyes to confirm perceptibility. With eyes open, the direction was adjusted to ensure consistent delivery to the eye corner. The airpuff was then presented two additional times while blink responses were monitored for characteristic blink-related peaks on electrooculography, indicating a successful response. Calibration was repeated as necessary until proper delivery was confirmed.

After the extinction phase, participants rated the airpuff (US) on valence and arousal after a single presentation, using the following questions: “How pleasant is the feeling that this airpuff evokes in you?” and “How intense or arousing is the feeling that this airpuff evokes in you?”.

SCR recording and preprocessing. SCRs were acquired via the GSR module (Brain Products GmbH, Germany), with electrodes placed on the lower thenar and hypothenar regions of the left hand. A conversion factor of 25 was applied online to transform the signal from microvolts into microsiemens. To minimize low-frequency drift and high-frequency noise, the data were filtered with a 0.05 Hz high-pass and a 5 Hz low-pass filter. The continuous signal was subsequently divided into epochs extending from 2,000 ms pre-CS to 6,000 ms post-CS, allowing analysis of both baseline and stimulus-related activity. The 2-second interval preceding CS onset was defined as the baseline period.

All epochs were manually inspected to identify artifacts caused by movement or excessive noise, and those containing such artifacts were excluded from further analysis. To correct for baseline activity, the mean signal during the baseline period was subtracted from the peak amplitude following CS onset (Kuhn et al., 2022). Any resulting negative values were set to zero. The baseline-corrected SCR amplitudes were then used in subsequent statistical analyses.

ECG recording and preprocessing. Using a BIP2AUX adapter, ECG signals were recorded with electrodes placed bilaterally at the midclavicular-subclavicular line intersections. To enhance R-peak detection, the signal underwent filtering with a 0.5 - 100 Hz band-pass filter and a 50 Hz notch filter. The data were segmented into epochs ranging from 3 s before to 12 s after CS onset, with time 0 corresponding to CS onset. R-peak detection was performed automatically using HRVTool (Vollmer, 2019), with manual correction applied when necessary. Subsequently, R-peaks were converted into interbeat intervals (RR intervals) and then transformed into instantaneous heart rate (HR) values, expressed in beats per minute.

For analysis purposes, HR data were interpolated into 0.5-second bins covering the full 15-second epoch, using linear interpolation based on the timing of R-peaks. RR intervals shorter than 273 ms (corresponding to heart rates above 220 beats per minute) or longer than 1.5 seconds (below 40 beats per minute) were excluded as physiologically implausible. Trials were discarded if more than 10 consecutive bins contained missing values (artifacts). For the remaining trials, linear interpolation was used to generate a continuous beats per minute time series, enabling time-resolved analysis of HR dynamics across the epoch. To correct for baseline, the mean heart rate (bpm) in the -2 to 0 s pre-stimulus interval was subtracted from each data point, with calculations performed separately for the different CS types. Previous studies have shown that HR responses to CS+ typically involve an initial deceleration reflecting basic stimulus registration, followed by acceleration and then a later deceleration (Battaglia et al., 2022). Based on this, and after baseline correction, HR data were averaged over two post-stimulus intervals: 1.0 - 3.0 seconds (early phase) and 3.0 - 5.5 seconds (late phase).

EEG recording and preprocessing. EEG recordings were obtained using a 64-channel ActiCHamp amplifier (Brain Products GmbH, Germany). Electrodes were placed according to the 10-20 system, with Cz as the online reference and Fpz as ground. Electrode impedances were maintained under 25 kΩ, and signals were digitized at 1000 Hz.

EEG preprocessing was carried out with EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts. Channels identified as noisy upon visual inspection were corrected using spherical spline interpolation. Data were re-referenced to the common average, with the Cz signal reconstructed to retain its original information. A two-stage filtering procedure was applied. First, a 1 Hz high-pass filter was applied to the data in preparation for independent component analysis (ICA), which was carried out using the AMICA algorithm (Palmer et al., 2019) to identify independent components reflecting artifacts such as eye movements, muscle activity, and other non-neural sources. The ICA weights were then transferred to the main EEG dataset, which had been filtered using a band-pass finite impulse response filter between 0.1 Hz and 45 Hz to remove slow drifts and high-frequency noise while preserving task-related neural signals. To reduce computational demands, both datasets were downsampled to a rate of 250 Hz. Following ICA, artifact-related components (e.g., eye blinks, muscle activity) were visually identified and removed. The continuous EEG data were then divided into epochs spanning -3000 ms to 7000 ms relative to the onset of the conditioned stimulus (CS; 0 ms). Epochs containing residual artifacts were manually inspected and excluded from further analysis.

For ERP analyses, signals were re-referenced to the linked mastoids (TP9/TP10). Baseline correction performed with the 500 ms interval before CS onset. In addition, automatic artifact rejection was conducted using a ±100 µV threshold, applied after temporarily low-pass filtering the data at 30 Hz to reduce the influence of high-frequency noise. Only epochs free of residual artifacts (as identified through automatic rejection) were retained for averaging. ERPs were computed separately for each participant and condition using ERPLAB. These averaged waveforms were used in subsequent statistical comparisons across conditions.

Time-frequency representations of oscillatory power were computed using Morlet wavelet convolution in Fieldtrip toolbox (Oostenveld et al., 2011). The time-frequency analysis was conducted on preprocessed single-trial data from 1 to 45 Hz in 1-Hz steps, using Morlet wavelets with cycle numbers varying from 3 to 12 across 45 logarithmically spaced steps, separately for each participant and CS type. Analyses focused on trials corresponding to the CSs (CS+ and CS-). For each condition, time-frequency decomposition was performed at each electrode with a temporal resolution of 20 ms across a window from -3000 ms to 7000 ms relative to CS onset. A relative change method was used for baseline correction, referencing the -600 to -100 ms interval preceding stimulus onset. For each participant, single-subject power spectra were averaged across trials, and these participant-level time-frequency representations were then entered into group-level statistical analyses.

Pupil size and eye blink recording and preprocessing. Pupil diameter was recorded using a head mounted Pupil Labs eye tracker at a sampling rate of 120 Hz. Data from the left eye were used for all analyses. Data preprocessing was carried out in R using the gazeR package (Geller et al., 2020). Trials were epoched from 3,000 ms before CS onset to 3,000 ms after US onset. A confidence score, provided by the eye-tracking system, reflects the estimated reliability of pupil size measurements, taking into account factors such as image clarity, pupil visibility, and tracking stability. Scores below 0.8 typically indicate conditions such as blinks, occlusion, or poor lighting, which compromise data reliability, and were thus excluded from analysis. Blink data were extracted from the eye-tracking output. Additionally, to identify blink-related artifacts, pupil dilation speed was calculated, and samples exceeding a trial-specific median absolute deviation threshold were marked as missing. Blink intervals were extended by 100 ms before and after each blink. Trials with more than 50% missing data were excluded, and participants with fewer than 10 valid trials per CS type were removed. Remaining data were linearly interpolated, smoothed using a moving average filter, baseline-corrected using the -1,000 to 0 ms pre-CS interval, and downsampled into 100 ms bins for analysis. Blink data were aligned to the same trial structure as the pupil data. Each time point was assigned a binary value (0 = eye open, 1 = blink), and data were downsampled into 100 ms bins spanning -3,000 to +6,000 ms relative to CS onset. For each time bin, the proportion of blink-labeled samples was calculated to produce a time-resolved blink profile for each trial. These data were used to evaluate blink rate dynamics across conditions.

Debriefing interview. After the conditioning experiment, participants underwent a debriefing interview. The interview was structured hierarchically, beginning with a general question "What did you see, feel?" and then "Was there anything else?" It then progressed to the more specific questions regarding contingency awareness "Were the shapes in any way predictive of the US (images, shocks, or airpuffs)?". If participants reported noticing a contingency, they were asked to indicate the block in which they became aware of it and estimate the approximate point within that block at which contingency awareness was acquired, using a scale from 1 to 100% (e.g., 30% into Block 2). At the end of the interview, the experimenter provided a full explanation of the CS-US contingency.

3.2.3 Contingency awareness criteria

Contingency awareness was assessed based on participants’ responses after each acquisition block. Specifically, after each block, participants answered contingency awareness questions and rated their US expectancy for both CS+ and CS-. A participant was classified as “aware” for that block if two criteria were met: (i) they gave a positive response to the contingency awareness questions and correctly identified the CS+, and (ii) they reported greater US expectancy for CS+ compared to CS-. Otherwise, they were classified as “unaware” for that block.

To improve the accuracy of this classification, two correction procedures were applied. First, participants’ retrospective reports from the post-experiment interview were reviewed. If their reported time of gaining awareness conflicted with the initial classification, we re-evaluated the block-wise awareness status using the pattern of US expectancy ratings as the final criterion (i.e., if participants reported greater US expectancy for CS+ relative to CS-, the block was reclassified as “aware”; otherwise, as “unaware”). Second, we assumed that once awareness was acquired, it would persist. Therefore, if a participant met the awareness criteria in one block but failed to do so in a subsequent block, the later result was treated as a false negative, and the participant was still classified as “aware”.

For all analyses, except those specifically examining the relationship between contingency awareness and conditioned responses, only blocks in which participants were deemed as aware of the CS-US contingency were included. As a result, different participants contributed data from different numbers of acquisition blocks depending on when they reached awareness. For example, a participant who became aware in Block 1 contributed data from all five acquisition blocks, whereas a participant who reached awareness in Block 5 contributed data from that block only. Importantly, most participants acquired contingency awareness early in the task. Among the 115 valid participants, only two remained unaware throughout all five acquisition blocks. Of the remaining 113 participants, 111 (98%) were classified as aware by Block 2 and therefore contributed data from four or five acquisition blocks. It should also be noted that the block-wise awareness assessment likely has introduced a prompting effect, which could have facilitated earlier awareness in some participants.

3.2.4 Statistical analyses

For self-reported affective ratings and physiological measures (SCRs and HR), two repeated-measures ANOVAs were conducted for each dependent variable. The first ANOVA included CS type (CS+, CS-) and Block (6 levels for affective ratings, that included Baseline, and 5 levels for physiological measures) as within-subject factors, and US modality (airpuff, image, shock) as a between-subject factor. For this analysis, eligible participants were those aware of the CS-US contingency in all five acquisition blocks and who had at least one usable trial for each CS type within every block. Despite this moderate reduction in sample size, the analysis allowed for a focused examination of how CRs evolved across learning blocks. In the second ANOVA, CS type was treated as a within-subject factor and US modality as a between-subject factor, with responses averaged separately for each CS type and US modality. This complementary analysis retained a larger sample and examined overall group differences in CRs, independent of the learning block.

ERP analyses focused on two components: the late positive potential (LPP) and the stimulus-preceding negativity (SPN). LPP amplitude was averaged within the 400-1000 ms window at the POz electrode, and SPN amplitude was averaged within the 3000-5950 ms window at Cz. These time windows were determined with reference to earlier studies (Bacigalupo & Luck, 2018; Wiemer et al., 2021) and visual inspection of the flattened grand average across participants and CS types, which revealed the most pronounced component activity within these intervals (Bowman et al., 2020). Each component was analyzed using a mixed-design ANOVA, with CS type (CS+, CS-) specified as the within-subject factor and US modality (airpuff, image, shock) as the between-subject factor.

Time-frequency analyses were performed on alpha-band power (9-15 Hz). Power values were baseline-corrected and expressed as percentage change from baseline. Mean alpha power was extracted from the 1.0-5.9 s window at the C4 and POz electrodes. We chose the C4 electrode for one of the analyses because the electrical shock was applied to the left wrist, and C4 corresponds to the right somatosensory cortex. In addition, C3, the homologous counterpart of C4 in the left hemisphere, was included as a control site to examine potential hemispheric differences or generalization of alpha suppression. We expected that the alpha or mu rhythm - the primary somatosensory rhythm (Gundlach et al., 2017; Pfurtscheller & Lopes da Silva, 1999) - would show a suppression at C4 in response to the CS, similar to the response to the shock itself, while this effect would not be observed in the other US modality groups. POz was selected on the basis of previous research (Bacigalupo & Luck, 2022; Panitz et al., 2019) where alpha suppression was present in response to CS+ in the visual cortex, sometimes even if the CS is auditory (Farkas et al., 2024), and also on the basis of grand average curves averaged across participants and conditions showing the maximal alpha suppression around POz channel. For each electrode site, a separate mixed-design ANOVA was performed, with CS type (CS+, CS-) as the within-subject factor and US modality (airpuff, image, shock) as the between-subject factor.

For the airpuff group, additional analyses were conducted on pupil diameter and blink rate. Pupil data were baseline-corrected and averaged within the 4.0-5.9 s interval following CS onset. Blink rate was calculated as the proportion of samples marked as blinks within the 5.5-5.9 s time window. For each measure, two repeated-measures ANOVAs were conducted: the first included CS type (CS+, CS-) and Block (1-5) as within-subject factors and was limited to participants deemed aware of the CS-US contingency in all five acquisition blocks and had at least one valid trial for both CS types in each block; the second included CS type as a within-subject factor, with responses averaged across blocks and all available participants included.

To assess the influence of contingency awareness on CRs, additional analyses were conducted using data from Block 1 in the aversive image group only. This restriction served two purposes: to allow comparison with a previous online version of the task, which included only self-report measures (Li et al., 2025), and to minimize the potential influence of the contingency awareness questions presented between blocks on responses in Block 2. EEG data were excluded from these analyses due to the lower signal-to-noise ratio and the need for more trials than a single block could provide to ensure adequate data quality. For each measure (self-reported affective ratings, SCR, and HR), a repeated-measures ANOVA was performed with CS type (CS+ vs. CS-) as a within-subject factor and awareness group (unaware vs. aware) as a between-subject factor.

Effect sizes for ANOVAs are reported as partial eta squared (η^2_p). Post hoc pairwise comparisons were conducted where appropriate, with effect sizes reported as Cohen's *d*. All statistical tests were two-tailed with a significance threshold of $p < 0.05$. All statistical analyses were conducted using R.

3.3 Results

3.3.1 Affective self-report ratings

We investigated the influence of US modality on evaluative conditioned responses - fear, arousal, and valence in ANOVAs that included US modality (airpuff, image, shock) as a between-subjects factor, and CS type (CS+, CS-) and Block (Baseline, A1 through A5) as within-subjects factors. A significant main effect of CS type was observed across all three affective dimensions, indicating that participants consistently rated the CS+ as more fear-inducing, arousing, and unpleasant than the CS- (see **Table 7 and Figures 7, 8, and 9**). A significant main effect of Block was observed for fear, arousal, and valence ratings, suggesting that conditioned affective responses evolved systematically over the course of learning. In contrast, no significant main effect of US modality was observed for any of the affective dimensions. Significant interactions were observed between US modality and CS type for valence ratings, between US modality and Block for fear ratings, and between CS type and Block across all three measures. Additionally, a significant three-way interaction (US modality × CS type × Block) also emerged for valence, suggesting that the pattern of evaluative learning differed across modalities over time. Full statistical results are presented in **Table 7**.

Table 7. Results of ANOVAs examining conditioned affective responses across blocks, with Block (Baseline, A1 to A5) and CS type (CS+ vs. CS-) as within-subjects factors, and US modality (airpuff, image, shock) as a between-subjects factor (*n* = 88).

Dependent variable	Test statistics	US modality	CS type	Block	US modality x CS type	US modality x Block	CS type x Block	US modality x CS type x Block
Fear	*F*	0.69	94.19	10.9	2.01	2.28	18.64	1.73
	p	0.503	**<0.001***	**<0.001***	0.140	**0.039***	**<0.001***	0.114
	η^2_p	0.02	0.53	0.11	0.05	0.05	0.18	0.04
	df	2, 85	1, 85	2.88, 245.05	2, 85	5.77, 245.05	3.04, 258.47	6.08, 258.47
Arousal	*F*	0.63	36.28	3.65	0.06	1.83	12.09	0.68
	p	0.532	**<0.001***	**0.009***	0.943	0.082	**<0.001***	0.676
	η^2_p	0.02	0.30	0.040	<0.01	0.04	0.13	0.02
	df	2, 85	1, 85	3.51, 298.31	2, 85	7.02, 298.31	3.14, 266.51	6.27, 266.51
Valence	*F*	1.02	73.71	10.14	6.16	1.18	9.88	2.15
	p	0.365	**<0.001***	**<0.001***	**0.003***	0.314	**<0.001***	**0.046***
	η^2_p	0.02	0.46	0.11	0.13	0.03	0.10	0.05
	df	2, 85	1, 85	3.70, 314.72	2, 85	7.41, 314.72	3.09, 262.47	6.18, 262.47

Notes. A1-A5 refer to ratings collected after Blocks 1 through 5, respectively. **p* < .05.

To clarify the temporal dynamics of conditioned affective responding, we conducted post hoc comparisons between CS+ and CS- at each block within each US modality.

***Fear ratings**.* At baseline, no significant difference between CS+ and CS- was observed in the airpuff group, *t*(85) = 1.22, *p* = 0.227, *d* = 0.23, or the shock group, *t*(85) = 0.17, *p* = .869, *d* = 0.03. In contrast, a moderate difference was present in the image group (*t*(85) = 2.30, *p* = .024, *d* = 0.40), suggesting a slight pre-existing bias. Beginning in Block 1, fear ratings for CS+ were significantly higher than CS- in all three US modality groups, with large effect sizes: airpuff, *t*(85) = 4.98, *p* < .001, *d* = 0.88; image, *t*(85) = 5.57, *p* < 0.001, *d* = 1.15; and shock, *t*(85) = 6.89, *p* < 0.001, *d* = 1.20. This differentiation remained robust in the image and shock groups through Block 5, as shown by the significant CS+ vs. CS- differences specifically in Block 5 (image: *t*(85) = 5.05, *p* < .001, *d* = 0.85; shock: *t*(85) = 4.62, *p* < .001, *d* = 0.83). In the airpuff group, however, the effect declined over time and was only marginally significant in Block 5 (*t*(85) = 1.99, *p* = .049, *d* = 0.43). (see **Figure 7**).

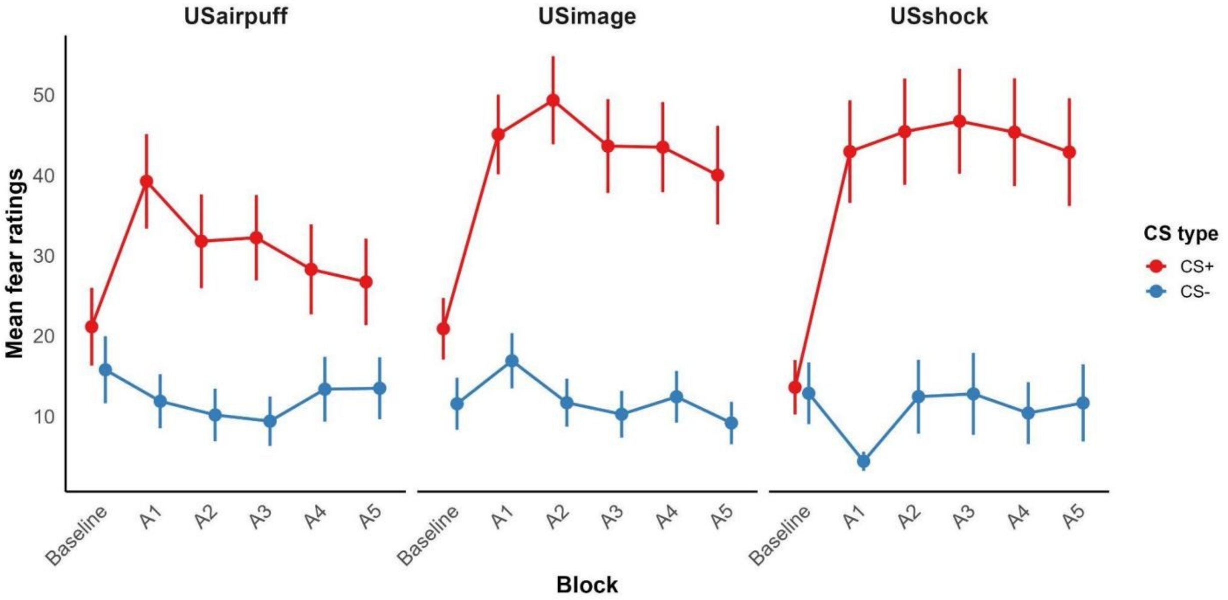


Figure 7. Mean fear ratings for CS+ and CS- across acquisition blocks, shown separately for the airpuff (*n* = 28), image (*n* = 33), and shock (*n* = 27) US modality groups. Error bars represent ±1 standard error (SE). A1-A5 refer to ratings collected after Blocks 1 to 5, respectively; the same labeling is used in **Figures 8 and 9**.

***Arousal ratings**.* At baseline, no significant differences in arousal ratings between CS+ and CS- were found in any of the US modality groups: airpuff, *t*(85) = -0.59, *p* = 0.557, *d* = -0.11; image, *t*(85) = 0.17, *p* = 0.867, *d* = 0.03; shock, *t*(85) = -0.92, *p* = 0.359, *d* = -0.18. Starting in Block 1, arousal ratings were consistently higher for CS+ than for CS- across all modalities, with moderate to large effect sizes: airpuff, *t*(85) = 3.85, *p* < 0.001, *d* = 0.63; image, *t*(85) = 4.44, *p* < 0.001, *d* =

0.82; and shock, *t*(85) = 3.32, *p* = 0.001, *d* = 0.71. Over time, arousal-based differentiation between CS+ and CS- remained most robust in the shock group, with a significant effect still evident in Block 5 (*t*(85) = 3.31, *p* = .001, *d* = 0.77). In contrast, the effect weakened in both the image (*t*(85) = 2.05, *p* = .043, *d* = 0.33) and airpuff (*t*(85) = 2.53, *p* = .013, *d* = 0.45) groups by Block 5, indicating a progressive reduction in conditioned arousal responses across the acquisition phase (see **Figure 8**).

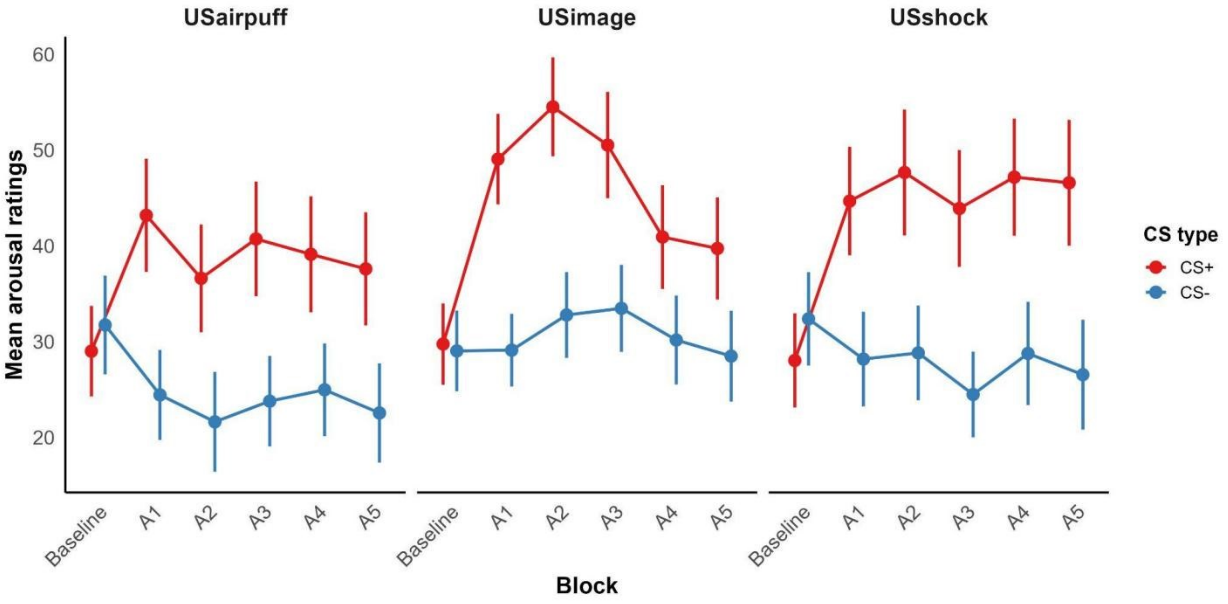


Figure 8. Mean arousal ratings for CS+ and CS- across acquisition blocks, shown separately for the airpuff (*n* = 28), image (*n* = 33), and shock (*n* = 27) US modality groups.

Valence ratings. At baseline, no significant differences in valence ratings between CS+ and CS- were observed in the airpuff (*t*(85) = 0.16, *p* = 0.875, *d* = 0.04) or shock (*t*(85) = -0.90, *p* = 0.369, *d* = -0.20) groups. However, a moderate and statistically significant difference was already present in the image group (*t*(85) = -2.90, *p* = 0.005, *d* = -0.40), indicating that CS+ was initially perceived as more negative than CS- prior to conditioning. In Block 1, participants in all three groups rated the CS+ as significantly more negative than the CS-: shock, *t*(85) = -4.88, *p* < .001, *d* = -1.00; image, *t*(85) = -4.31, *p* < .001, *d* = -0.74; airpuff, *t*(85) = -3.50, *p* = .001, *d* = -0.63. From Blocks 2 to 5, the shock and image groups continued to show strong and consistent CS+ vs. CS- differences (image: *t*(85) = -4.26, *p* < 0.001, *d* = -0.72; shock: *t*(85) = -4.97, *p* < 0.001, *d* = -0.93). In contrast, the effect in the airpuff group declined over time. A moderate difference remained in block 2 (*t*(85) = -2.05, *p* = 0.043, *d* = -0.38), but from block 3 onward, no significant differences were observed (e.g., block 3: *t*(85) = -1.04, *p* = 0.302, *d* = -0.22), suggesting that valence-based conditioned responding in the airpuff group was short-lived (see **Figure 9**).

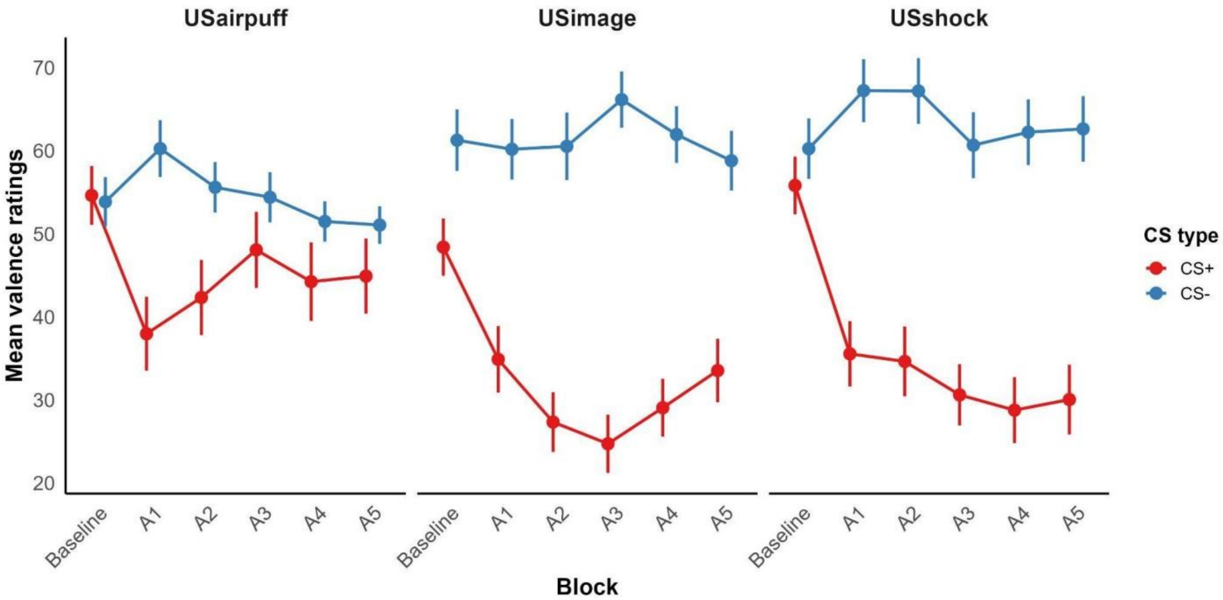


Figure 9. Mean valence ratings for CS+ and CS- across acquisition blocks, shown separately for the airpuff (*n* = 28), image (*n* = 33), and shock (*n* = 27) US modality groups.

To further examine CS discrimination independent of temporal variation, we conducted 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVAs, excluding the Block factor. These analyses revealed a consistent main effect of CS type across all affective measures, with no significant main effects of US modality. Significant CS type × US modality interactions were found for fear and valence ratings, but not for arousal. See **Table 8 and Figures 10-12** for details.

Post hoc comparisons indicated that CS+ elicited significantly stronger fear responses than CS- across all US modalities: airpuff, *t*(110) = 3.59, *p* = .0005, *d* = 0.64; image, *t*(110) = 8.16, *p* < .001, *d* = 1.18; shock, *t*(110) = 6.59, *p* < .001, *d* = 1.13. For valence ratings, CS+ was rated significantly more negative than CS- in the image and shock groups, but not in the airpuff group: airpuff, *t*(110) = -1.94, *p* = .055, *d* = -0.37; image, *t*(110) = -8.05, *p* < .001, *d* = -1.08; shock, *t*(110) = -6.26, *p* < .001, *d* = -1.17. For arousal, CS+ was rated as significantly more arousing than CS- across all US modalities: airpuff, *t*(110) = 3.49, *p* = .001, *d* = 0.60; image, *t*(110) = 4.84, *p* < .001, *d* = 0.67; shock, *t*(110) = 3.29, *p* = .001, *d* = 0.63.

Taken together, these findings demonstrate that US modality modulated various dimensions of conditioned affective responding. Specifically, both the shock and image modalities elicited stronger effects across all affective self-report measures, including fear, arousal, and valence, compared to the airpuff modality. Overall, the shock modality produced the most robust and consistent affective learning, indicating a particularly strong influence on affective conditioning.

Table 8. Results of ANOVAs examining overall conditioned affective responses, with CS type (CS+ vs. CS-) as a within-subjects factor, and US modality (airpuff, image, shock) as a between-subjects factor (*n* = 113).

Dependent variable	Source of variation	*F* (df_1, df_2)	*p*	η^2_p
Fear	CS type	106.51 (1, 110)	**<0.001***	0.49
	US modality	0.68 (2, 110)	0.509	0.01
	CS type × Modality	3.86 (2, 110)	**0.024***	0.07
Arousal	CS type	42.54 (1, 110)	**<0.001***	0.28
	US modality	0.37 (2, 110)	0.693	0.01
	CS type × Modality	0.07 (2, 110)	0.932	<0.01
Valence	CS type	82.44 (1, 110)	**<0.001***	0.43
	US modality	0.26 (2, 110)	0.771	<0.01
	CS type × Modality	7.89 (2, 110)	**<0.001***	0.13

Note. **p* < .05.

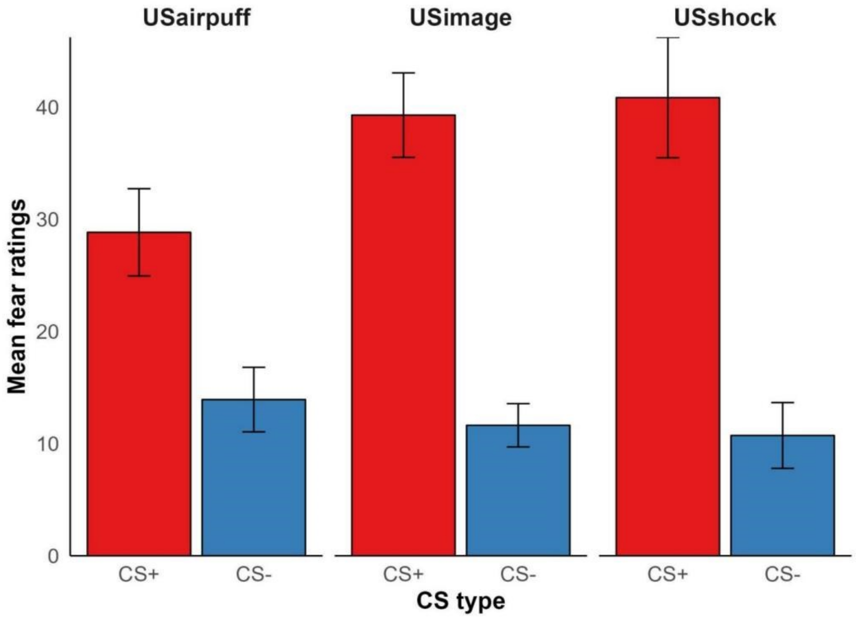


Figure 10. Mean fear ratings for CS+ and CS-, averaged across all blocks, shown separately for the image (*n* = 51), shock (*n* = 28), and airpuff (*n* = 34) US modality groups. Error bars represent ±1 standard error (SE).

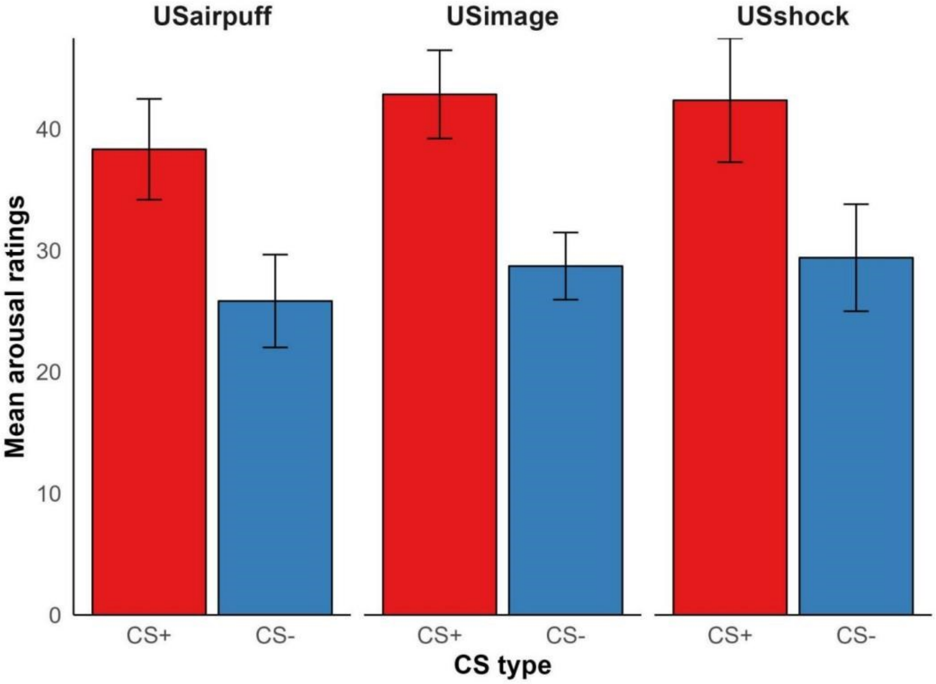


Figure 11. Mean arousal ratings for CS+ and CS-, averaged across all blocks, shown separately for the image (*n* = 51), shock (*n* = 28), and airpuff (*n* = 34) US modality groups.

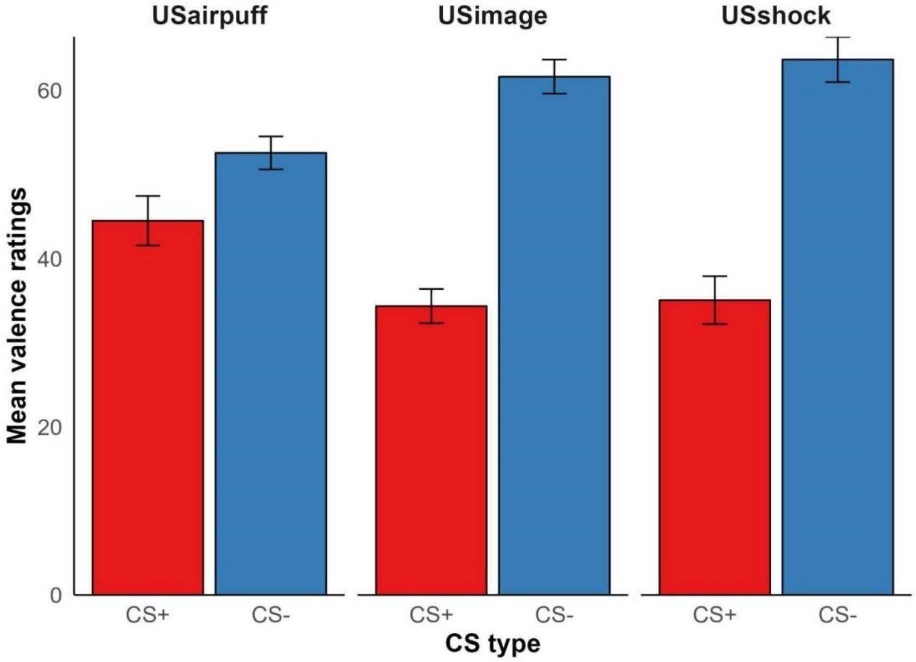


Figure 12. Mean valence ratings for CS+ and CS-, averaged across all blocks, shown separately for the image (*n* = 51), shock (*n* = 28), and airpuff (*n* = 34) US modality groups.

3.3.2 SCR

A 2 (CS type: CS+, CS-) × 5 (Block: 1-5) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on skin conductance responses (SCRs). The analysis revealed a significant main effect of CS type, *F*(1, 84) = 4.67, *p* = 0.034, η^2_p = 0.05, indicating that SCRs were generally higher for CS+ than for CS-. A significant main effect of Block was also observed, *F*(2.80, 235.11) = 4.37, *p* = 0.006, η^2_p = 0.05, reflecting changes in SCRs over time. Additionally, the US modality × Block interaction was significant, *F*(5.60, 235.11) = 3.93, *p* = 0.001, η^2_{p} = 0.09, suggesting that the temporal profile of SCRs differed by modality. No other main effects or interactions reached significance (*ps* > .05). The CS type × US modality interaction approached significance, *F*(2, 84) = 2.50, *p* = 0.088, η^2_p = 0.06. The main effect of US modality also trended toward significance, *F*(2, 84) = 2.37, *p* = 0.100, η^2_p = 0.05. Post hoc comparisons following the CS type × US modality interaction revealed that the difference between CS+ and CS- was significant only in the shock group, *t*(84) = 2.78, *p* = 0.007, *d* = 0.19. No significant differences were found in the airpuff group, *t*(84) = -0.34, *p* = 0.734, *d* = -0.05, or in the image group, *t*(84) = 1.25, *p* = 0.214, *d* = 0.09 (see **Figure 13**).

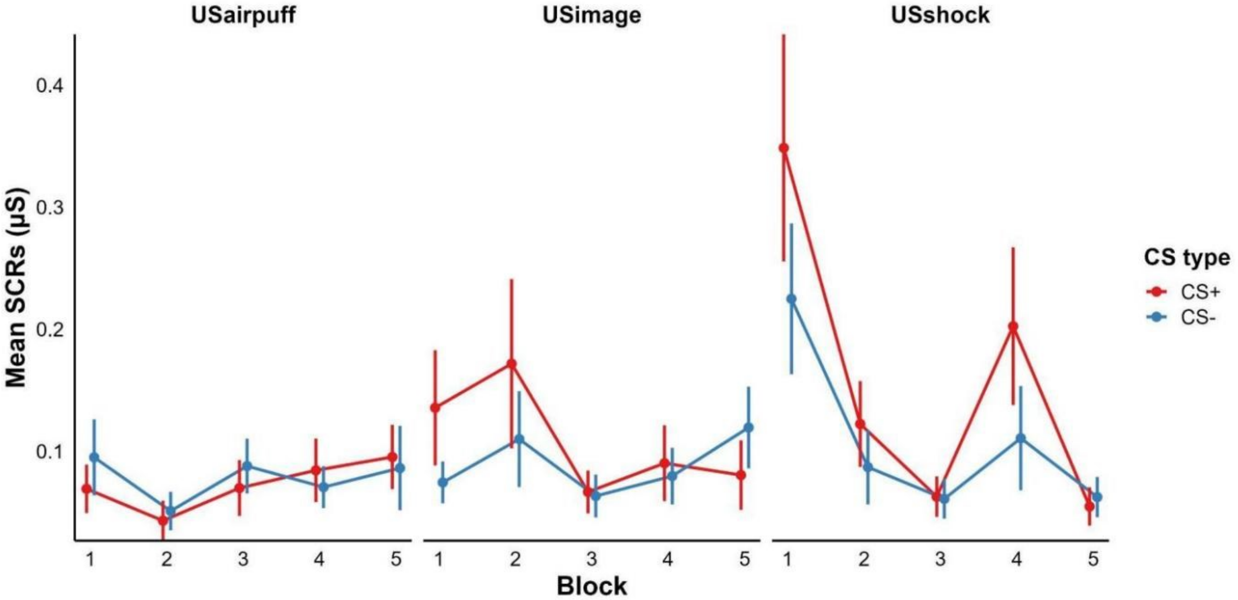


Figure 13. Mean skin conductance responses for CS+ and CS- across experimental blocks, shown separately for the image (*n* = 33), shock (*n* = 26), and airpuff (*n* = 28) US modality groups. Error bars represent ±1 standard error (SE).

To complement these findings, a 2 (CS type) × 3 (US modality) mixed-design ANOVA was conducted, collapsing across blocks. This analysis revealed a significant main effect of CS type, $F(1, 109) = 5.94$, $p = .016$, $\eta^2_p = 0.05$, again indicating overall higher SCRs to CS+ than to CS-. The main effect of US modality was not significant ($p = 0.349$), and the CS type × US modality interaction approached significance, $F(2, 109) = 2.39$, $p = .097$, $\eta^2_p = 0.04$. Exploratory post hoc comparisons revealed a significant CS+ > CS- difference only in the shock group, $t(109) = 2.92$, $p = .004$, $d = 0.46$. No significant differences were found in the airpuff ($p = .819$) or image ($p = .415$) groups (see **Figure 14**).

Collectively, these findings indicate that although SCRs were generally higher for CS+ than for CS-, this effect was largely specific to the shock modality.

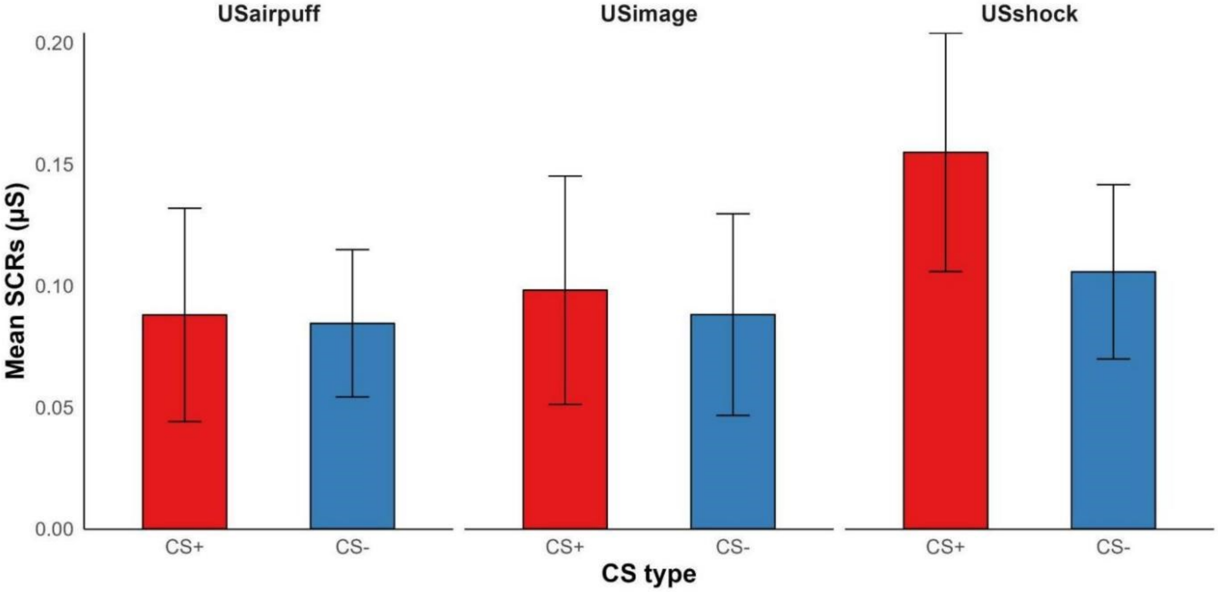


Figure 14. Mean skin conductance responses for CS+ and CS-, averaged across all blocks, shown separately for the image ($n = 51$), shock ($n = 27$), and airpuff ($n = 34$) US modality groups. Error bars represent ±1 standard error (SE).

3.3.3 Heart rate

Figure 15 shows the time course of baseline-corrected heart rate (HR) responses following CS onset, shown separately for the image ($n = 33$), shock ($n = 20$), and airpuff ($n = 22$) US modality groups.

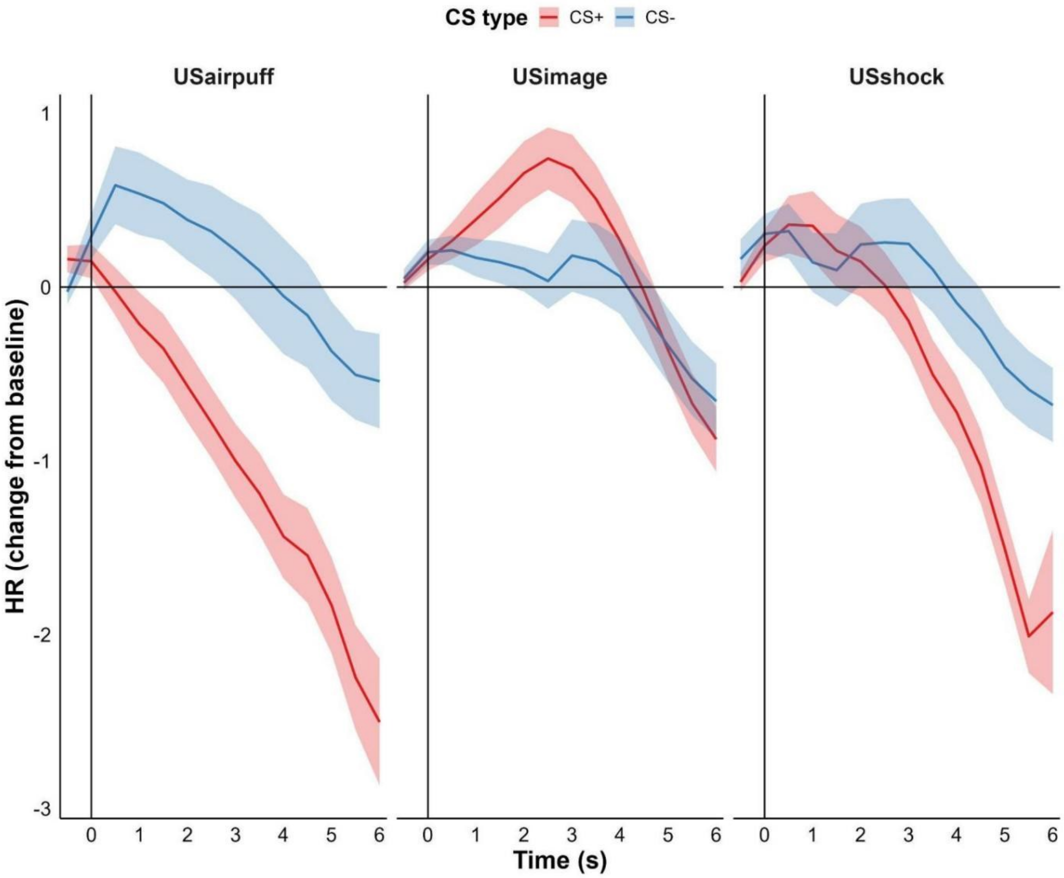


Figure 15. Time course of HR changes from baseline following CS onset, shown separately for the image (*n* = 33), shock (*n* = 20), and airpuff (*n* = 22) US modality groups. Shaded areas represent ±1 standard error (SE).

A 2 (CS type: CS+, CS-) × 5 (Block: 1-5) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on HR responses during the early phase (1.0-3.0 s post-CS onset). The analysis revealed a significant CS type × US modality interaction, $F(2, 72) = 14.82$, $p < 0.001$, η^2_p= 0.29. In addition, three main effects approached significance: for CS type, $F(1, 72) = 3.50$, $p = 0.065$, η^2_p= 0.046; for US modality, $F(2, 72) = 2.62$, $p = 0.079$, η^2_p= 0.068; and for Block, $F(3.82, 275.09) = 2.13$, $p = 0.081$, η^2_p= 0.029. No other main effects or interactions reached significance (*ps* > 0.05). Follow-up comparisons on the CS type × US modality interaction revealed modality-specific HR patterns. In the image group, CS+ elicited greater HR acceleration than CS-, $t(72) = 2.53$, $p = 0.014$, $d = 0.16$, reflecting a small effect size. In contrast, in the air puff group, CS+ elicited HR deceleration, whereas CS- elicited acceleration, with a significant difference, $t(72) = -4.95$, $p < 0.001$, $d = -0.42$. No significant difference was observed in the shock group, $t(72)$ = -

0.22, *p* = .826, *d* = -0.02. **Figure 16** shows mean HR changes during the early phase by CS type, US modality, and block.

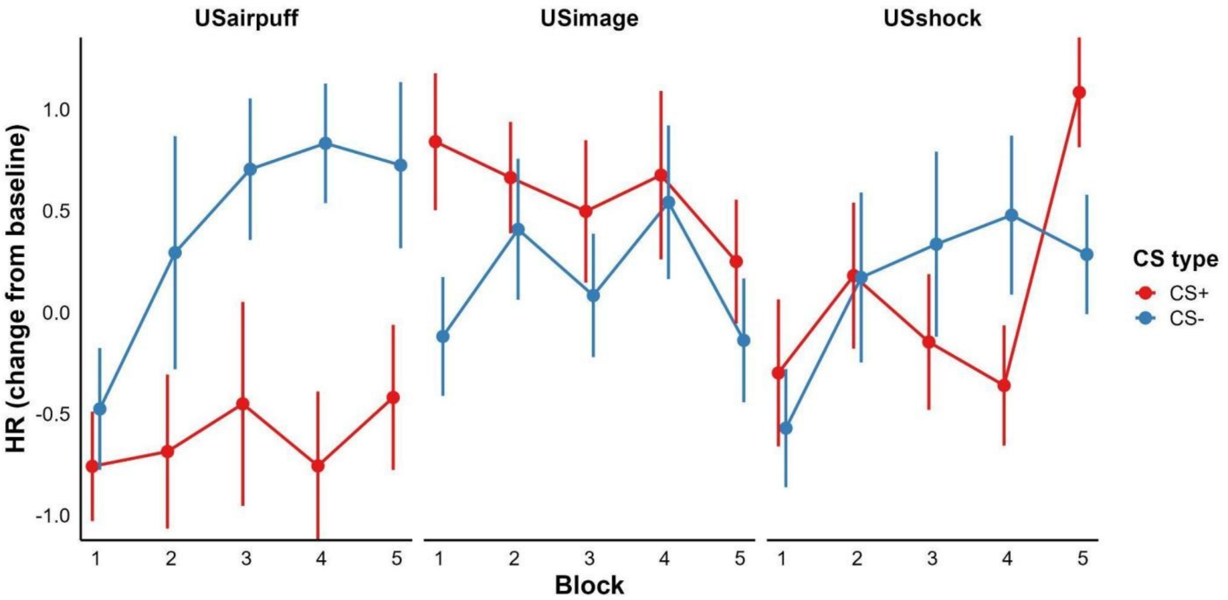


Figure 16. Mean HR changes from baseline during the early phase (1.0-3.0 s post-CS onset), for CS+ and CS- across blocks, shown separately by US modality: image (*n* = 33), shock (*n* = 20), and airpuff (*n* = 22). Error bars represent ±1 standard error of the mean (SE).

Another 2 (CS type: CS+, CS-) × 5 (Block: 1-5) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on HR responses during the late phase (3.0-5.5 s post-CS onset). The analysis revealed a significant main effect of CS type, *F*(1, 72) = 19.97, $p < 0.001$, $\eta^2_p = 0.22$, with CS+ trials eliciting greater HR deceleration (i.e., more negative change-from-baseline values) compared to CS- trials. A significant main effect of US modality was also observed, *F*(2, 72) = 6.97, *p* = 0.002, $\eta^2_p = 0.16$, indicating that overall HR responses varied across modalities. Critically, the CS type × US modality interaction was significant, *F*(2, 72) = 9.15, $p < 0.001$, $\eta^2_p = 0.20$. No other main effects or interactions reached significance (*ps* > 0.05). Follow-up comparisons on the CS type × US modality interaction revealed modality-specific patterns of HR deceleration. In the airpuff group, CS+ trials produced greater HR deceleration than CS-, *t*(72) = -5.09, $p < .001$, *d* = -0.50. A similar pattern was observed in the shock group, *t*(72) = -2.57, *p* = .012, *d* = -0.35, with CS+ again producing greater deceleration. In contrast, the image group showed no significant difference, *t*(72) = 0.43, *p* = .666, *d* = 0.03. **Figure 17** presents HR changes during the late phase.

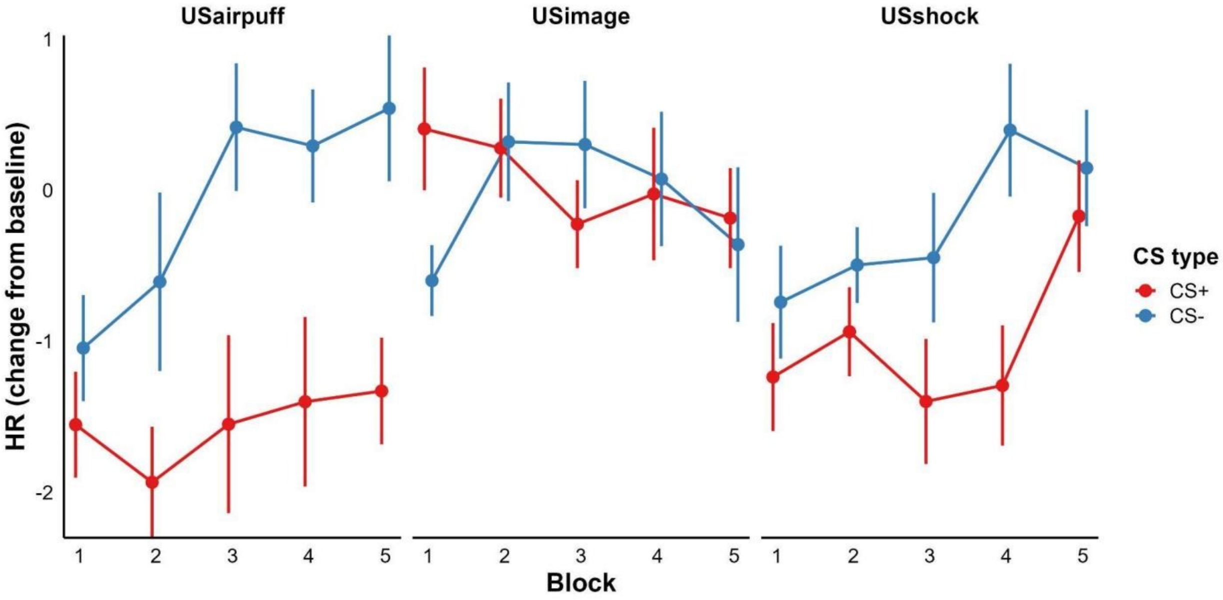


Figure 17. Mean HR changes from baseline during the late phase (3.0-5.5 s post-CS onset), for CS+ and CS- across blocks, shown separately by US modality: image (*n* = 33), shock (*n* = 20), and airpuff (*n* = 22). Error bars represent ±1 standard error (SE).

To complement the block-wise analysis during the early phase, a 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted, collapsing across blocks. The analysis revealed a significant CS type × US modality interaction, $F(2, 102) = 9.40$, $p < 0.001$, $\eta^2_p = 0.16$. A trend-level main effect of CS type was also observed, $F(1, 102) = 3.16$, $p = 0.079$, $\eta^2_p = 0.03$. The main effect of US modality did not reach significance, $F(2, 102) = 1.81$, $p = 0.169$, $\eta^2_p = 0.03$. Post-hoc comparisons highlighted differing HR patterns across US modalities. In the image group, CS+ elicited significantly stronger HR acceleration than CS-, $t(102) = 1.99$, $p = 0.0498$, $d = 0.35$. In contrast, in the airpuff group, CS+ produced HR deceleration while CS- produced acceleration, $t(102) = -3.96$, $p < 0.001$, $d = -0.60$. No significant difference was found in the shock group, $t(102) = -0.61$, $p = .544$, $d = -0.11$.

To complement the block-wise analysis during the late phase, a 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on HR change values, with data collapsed across blocks. The analysis revealed significant main effects of CS type, $F(1, 102) = 16.82$, $p < 0.001$, $\eta^2_p = 0.14$, and US modality, $F(2, 102) = 8.21$, $p < 0.001$, $\eta^2_p = 0.14$. The CS type × US modality interaction was also significant, $F(2, 102) = 4.09$, $p = 0.019$, $\eta^2_p = 0.07$, indicating that the extent of HR deceleration in response to CS+ versus CS- differed across US modalities. Post-hoc comparisons revealed the following modality-specific patterns. In the airpuff group, CS+ elicited significantly greater HR deceleration than CS-, $t(102) = -4.00$, $p < 0.001$, $d = -0.55$. A similar effect was observed in the shock group, with CS+ again producing stronger deceleration, $t(102) = -2.29$, $p = 0.024$, $d = -0.63$. In contrast, no significant difference was found in the image group ($p = 0.587$).

Taken together, these findings demonstrate that conditioned HR responses were modulated by US modality and varied over time. In the image group, CS+ elicited greater HR acceleration than CS- during the early phase, but no differences emerged during the late phase. In contrast, both the airpuff and shock groups exhibited stronger HR deceleration to CS+ than to CS- during the late phase, with no evidence of early-phase acceleration.

3.3.4 EEG measures

ERP components

SPN at Cz

A 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on SPN amplitudes at electrode Cz. The analysis revealed a significant main effect of CS type, $F(1, 106) = 15.76$, $p < 0.001$, $\eta^2_p = 0.07$, with CS+ eliciting larger SPNs than CS-. The main effect of US modality was not significant ($p = 0.660$). However, a marginally significant CS type × US modality interaction was observed, $F(2, 106) = 2.88$, $p = 0.060$, $\eta^2_p = 0.03$.

Post hoc comparisons revealed that the CS+ vs. CS- difference in SPN amplitude was significant in the shock group, $t(106) = -3.27$, $p = 0.001$, $d = -0.45$, and the airpuff group, $t(106) = -2.66$, $p = 0.009$, $d = -0.62$. In contrast, no significant difference was observed in the image group ($p = 0.545$).

Figure 18 displays the full ERP time course at Cz (-200 to 6000 ms), with the SPN window (3000-5950 ms) marked at Cz. The plot highlights the anticipatory differences between CS+ and CS-, particularly in the shock and airpuff groups. **Figure 19** displays the scalp topographies of SPN, averaged for CS+ and CS- in each US modality group.

LPP at POz

A separate 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA was conducted on LPP amplitudes at electrode POz. The analysis revealed a significant main effect of CS type, *F*(1, 106) = 10.07, *p* = 0.002, η^2_p= 0.02, with CS+ eliciting larger LPPs than CS-. The main effect of US modality approached significance, *F*(2, 106) = 2.52, *p* = 0.085, η^2_p= 0.04, while the CS type × US modality interaction was not significant (*p* = 0.168).

Exploratory post hoc comparisons showed a significant CS+ > CS- difference in LPP amplitude in the US shock group, *t*(106) = 2.98, *p* = .004, *d* = 0.52. No significant differences were observed in the US airpuff (*p* = .150) or image (*p* = .424) groups. **Figure 20** shows the ERP waveforms from 0 to 1000 ms, focusing on the LPP window (400-1000 ms) at POz. Differences between CS+ and CS- are most evident in the shock group. To illustrate the scalp distribution, **Figure 21** presents topographical maps of LPP, averaged for CS+ and CS- in each US modality group.

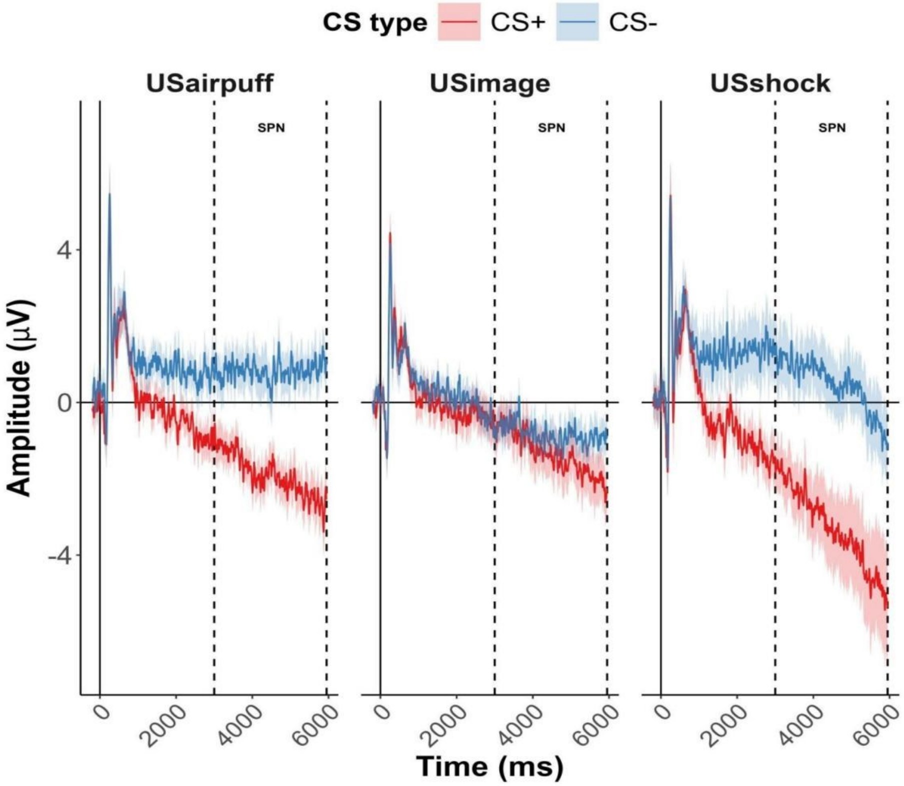


Figure 18. ERP waveforms at Cz (-200 to 6000 ms), shown separately for CS+ and CS- in each US modality group: image (*n* = 48), shock (*n* = 27), and airpuff (*n* = 34). Dashed vertical lines mark the SPN window (3000-5950 ms). Shaded ribbons represent ±1 SE.

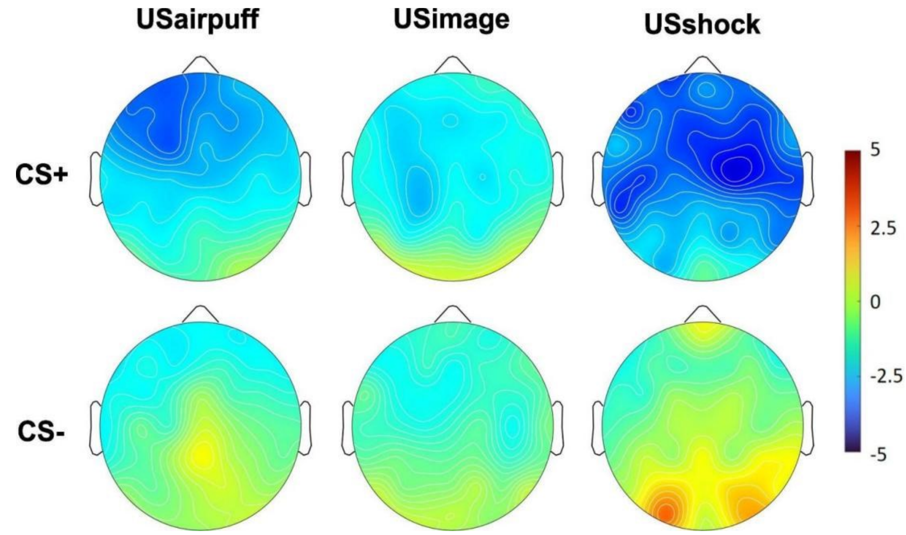


Figure 19. Topographical maps of SPN (3000-5950 ms) averaged for CS+ and CS- trials, shown separately for each US modality group: airpuff (*n* = 34), image (*n* = 48), and shock (*n* = 27). The color scale reflects ERP amplitude in µV.

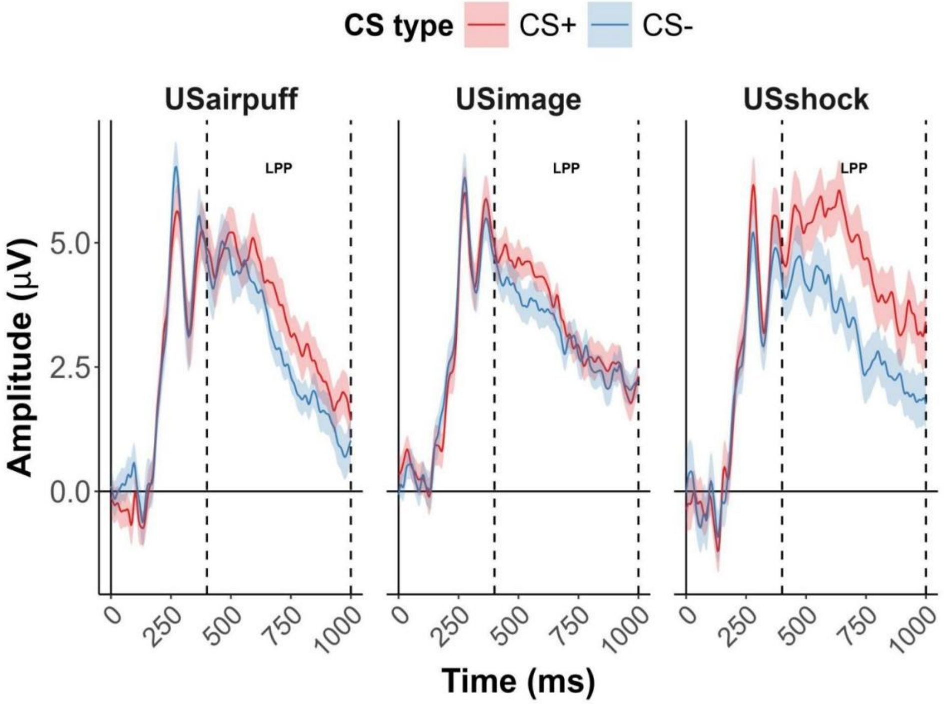


Figure 20. ERP waveforms at POz (0 to 1000 ms), shown separately for CS+ and CS- in each US modality group: image (*n* = 48), shock (*n* = 27), and airpuff (*n* = 34). Dashed vertical lines indicate the LPP window (400 to 1000 ms). Shaded ribbons represent ±1 SE.

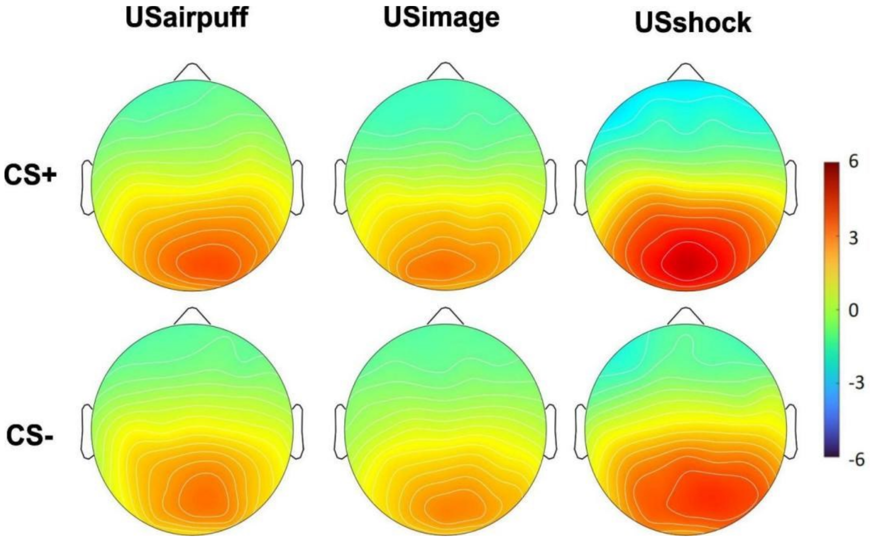


Figure 21. Topographical maps of LPP (400-1000 ms) averaged for CS+ and CS- trials, shown separately for each US modality group: airpuff (*n* = 34), image (*n* = 48), and shock (*n* = 27). The color scale reflects ERP amplitude in µV.

***Alpha power suppression*.** Time-frequency maps at C4, C3, and POz are shown in **Figures 22, 23, and 24**, respectively**. Figure 25** presents alpha power time courses at C4, C3, and POz across all US modality groups. **Figure 26** displays the topographical distribution of alpha power across the scalp.

C4 electrode

A 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA on alpha power at the C4 electrode revealed significant main effects of US modality, $F(2, 106) = 4.83$, $p = 0.010$, $\eta^2_p = 0.08$, and CS type, $F(1, 106) = 7.77$, $p = 0.006$, $\eta^2_p = 0.07$. The US modality × Condition interaction approached significance, $F(2, 106) = 2.76$, $p = 0.068$, $\eta^2_p = 0.05$.

Follow-up comparisons showed that alpha power was significantly lower for CS+ than CS- in the shock group, $t(106) = -2.92$, $p = 0.004$, $d = -0.68$. No significant differences were found in the airpuff ($p = 0.132$) or image ($p = 0.995$) groups, suggesting enhanced alpha suppression to CS+ specifically in the shock condition.

C3 electrode

A comparable 2 (CS type: CS+, CS-) × 3 (US modality: airpuff, image, shock) mixed-design ANOVA on alpha power at the C3 electrode revealed no significant main effects of US modality, $F(2, 106) = 0.55$, $p = 0.578$, $\eta^2_p = 0.01$, or CS type, $F(1, 106) = 0.01$, $p = 0.928$, $\eta^2_p < 0.01$. The interaction between US modality and CS type was also nonsignificant, $F(2, 106) = 0.33$, $p = 0.723$, $\eta^2_p = 0.01$. These results suggest that alpha activity at C3, a left-hemisphere homolog of C4, did not reliably differentiate CS+ from CS- across modalities.

POz electrode

A similar 2 × 3 ANOVA conducted at the POz electrode revealed no significant main effects of US modality, $F(2, 106) = 0.72$, $p = 0.488$, $\eta^2_p = 0.01$, or CS type, $F(1, 106) = 1.78$, $p = 0.185$, $\eta^2_p = 0.02$. The US modality × Condition interaction was also nonsignificant, $F(2, 106) = 0.16$, $p = 0.848$, $\eta^2_p < 0.01$. These findings indicate that alpha power at POz did not significantly vary by CS type or modality.

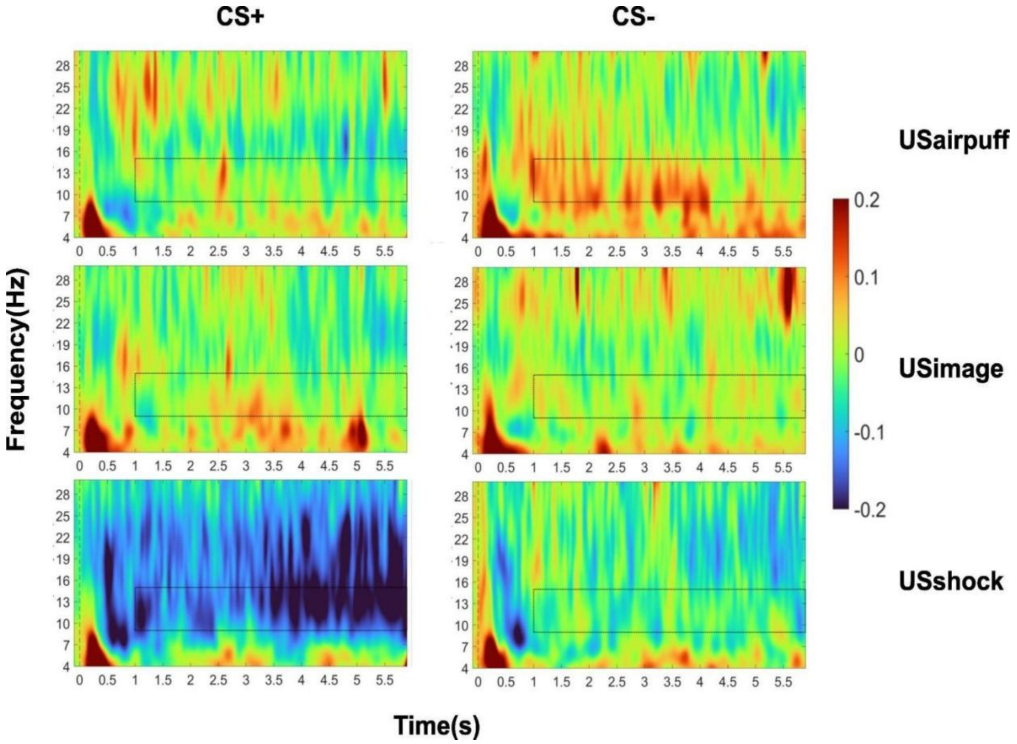


Figure 22. Time-frequency representations of baseline-normalized spectral power at channel C4, averaged across CS+ and CS- trials for each US modality group (image: n = 48; shock: n = 27; airpuff: n = 34). The dashed box indicates the alpha band (9-15 Hz) and time window (1.0-5.9 s). The color bar represents the relative change in power compared to baseline.

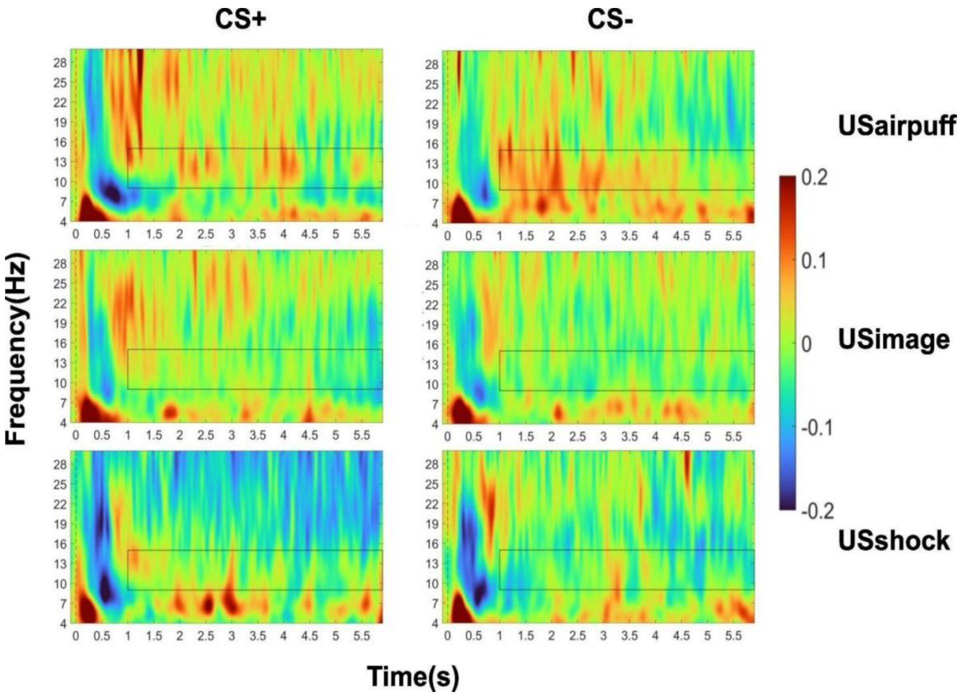


Figure 23. Time-frequency representations of baseline-normalized spectral power at channel C3, averaged across CS+ and CS- trials for each US modality group (image: *n* = 48; shock: *n* = 27; airpuff: *n* = 34).

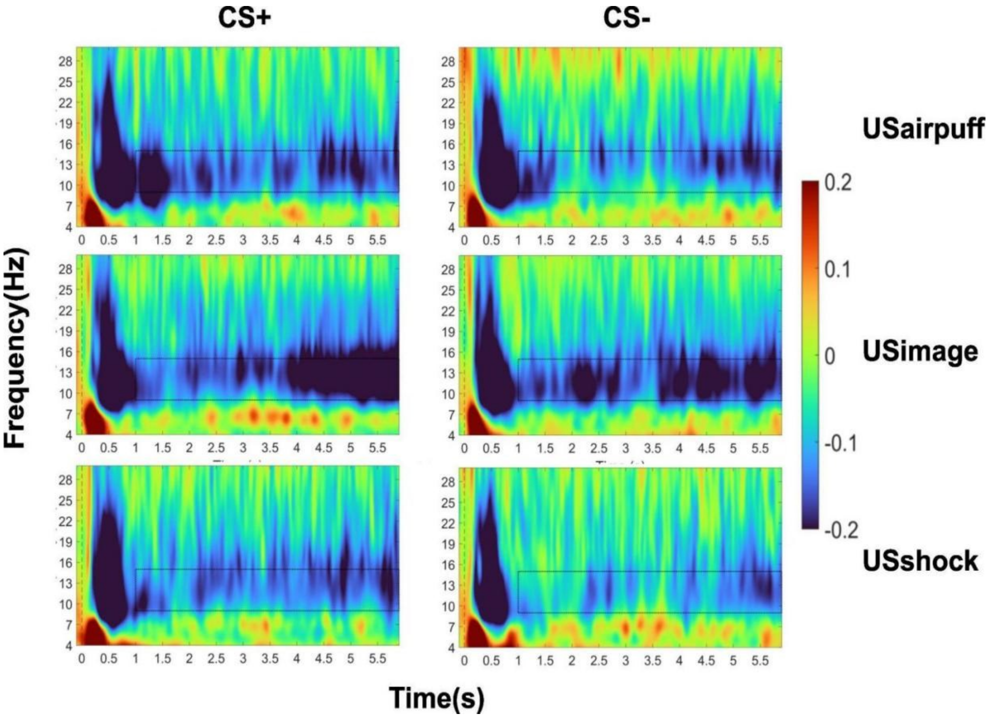


Figure 24. Time-frequency representations of baseline-normalized spectral power at channel POz, averaged across CS+ and CS- trials for each US modality group (image: *n* = 48; shock: *n* = 27; airpuff: *n* = 34).

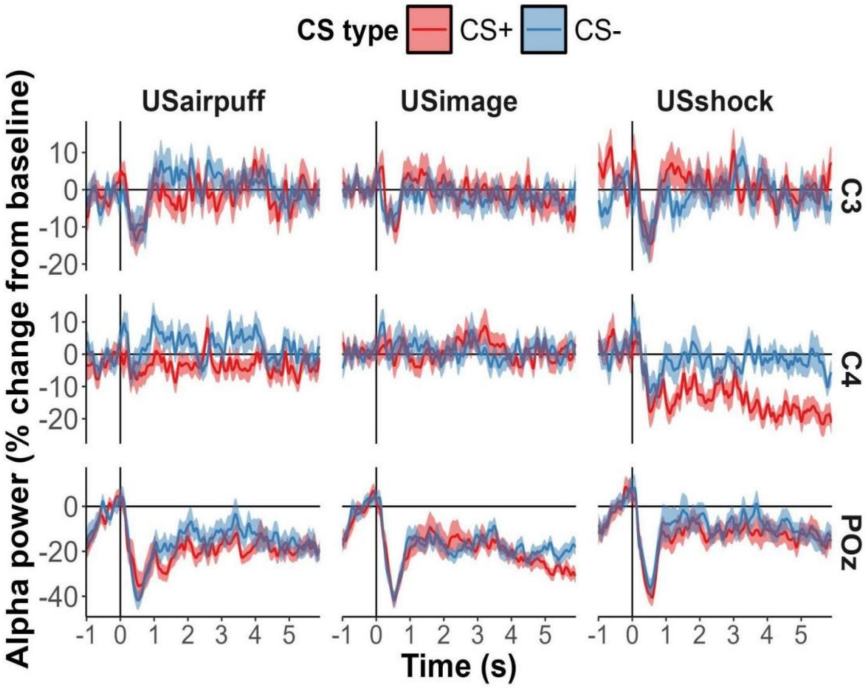


Figure 25. Alpha power time courses (% change from baseline) at electrodes C4, C3, and POz for CS+ and CS- conditions, shown separately for each US modality group (image: *n* = 48; shock: *n* = 27; airpuff: *n* = 34). Shaded areas indicate ±1 standard error (SE).

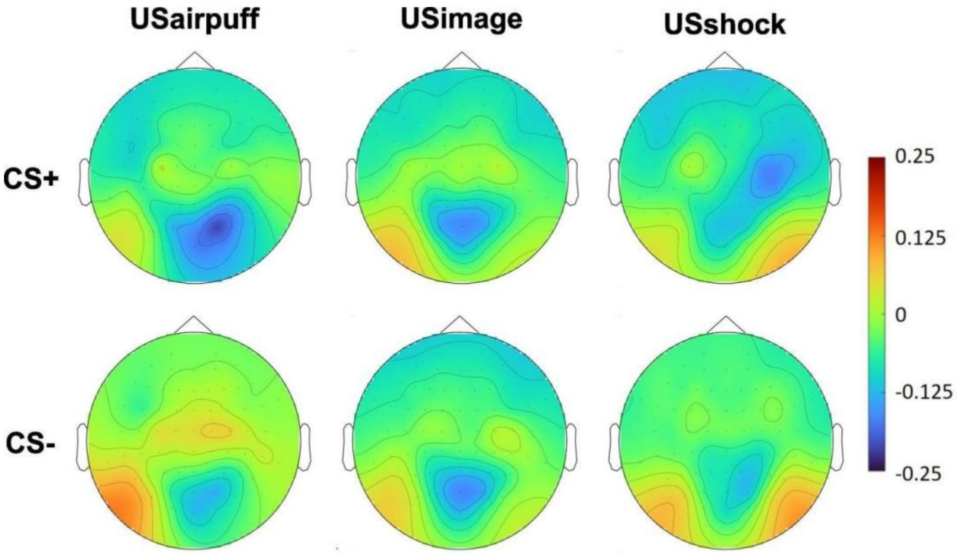


Figure 26. Topographical distribution of baseline-normalized alpha power (9-15 Hz) averaged over the 1.0-5.9 s time window across all EEG channels. Separate maps are shown for CS+ and CS- trials in each US modality group: airpuff (*n* = 34), image (*n* = 48), and shock (*n* = 27). The color bar represents the relative change in power compared to baseline.

3.3.5 Pupillometry and blink probability

Pupil diameter. In the airpuff group, to assess differences in pupil diameter, a 2 (CS type: CS+, CS-) × 5 (Block: 1-5) repeated-measures ANOVA was conducted. There were no significant main effects of CS type, $F(1, 15) = 0.39$, $p = 0.540$, $\eta^2_p = 0.03$, or Block, $F(2.95, 44.23) = 1.15$, $p = 0.338$, $\eta^2_p = 0.07$. The CS type × Block interaction was also not significant, $F(3.40, 50.93) = 1.73$, $p = 0.167$, $\eta^2_p = 0.10$. See **Figure 27** for the time course of pupil size in response to CS+ and CS- across five blocks.

To include more participants, we conducted an analysis using pupil diameter averaged across blocks as the dependent variable, with CS type (CS+, CS-) as the only within-subjects factor. The analysis revealed no significant effect of CS type, $F(1, 22) = 1.53$, $p = 0.230$, $\eta^2_p = 0.06$.

Taken together, these results indicate that in the airpuff group, pupil diameter did not significantly differ between CS+ and CS-.

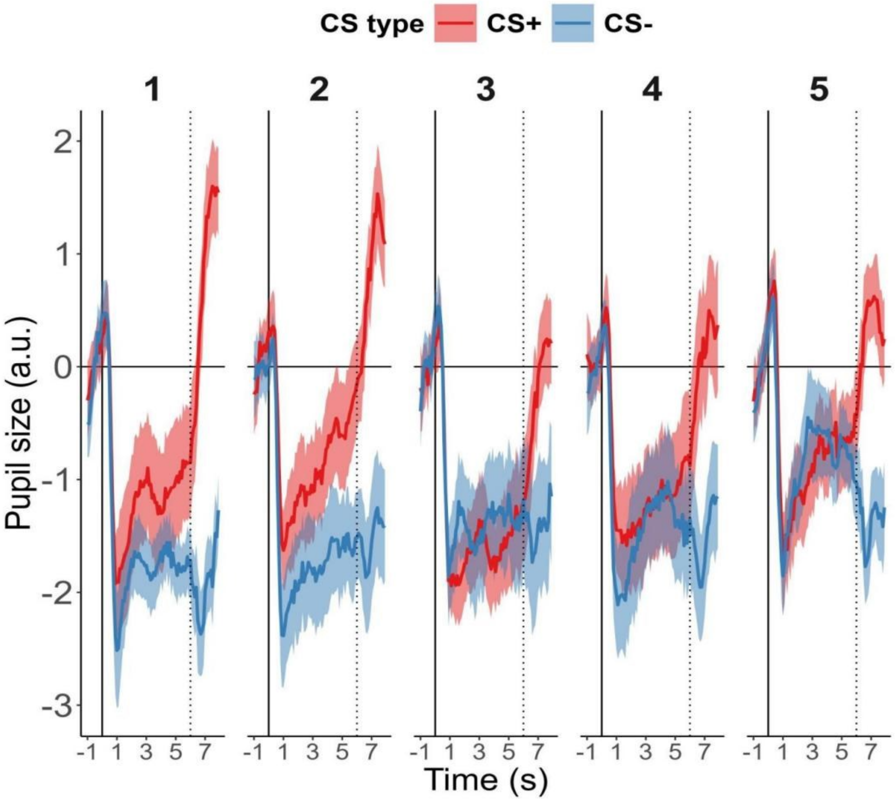


Figure 27. Time course of pupil size (in arbitrary units) in response to CS+ and CS- across five blocks in the airpuff group ($n = 16$). Pupil diameter was analyzed in the 4.0-5.9 s time window following CS onset. The dotted vertical line at 6 s indicates the onset of the unconditioned stimulus (US). Shaded areas indicate ±1 standard error (SE).

Blink probability. A 2 (CS type: CS+, CS-) × 5 (Block) repeated-measures ANOVA was conducted to assess differences in blink probability between CS types and across blocks. The analysis revealed no significant main effects of CS type, $F(1, 24) < 0.01$, $p = 0.974$, $\eta^2_p < 0.01$, or Block, $F(2.86, 68.56) = 1.20$, $p = 0.316$, $\eta^2_p = 0.05$. The interaction was also non-significant, $F(2.93, 70.36) = 1.72$, $p = 0.172$, $\eta^2_p = 0.07$ (see **Figure 28**).

To include a larger sample, we repeated the analysis with CS type as the only within-subjects factor, collapsing across blocks. This analysis again yielded no significant effect, $F(1, 31) = 0.12$, $p = 0.736$, $\eta^2_p < 0.01$.

In sum, blink probability did not significantly differ between CS+ and CS- trials in the airpuff group.

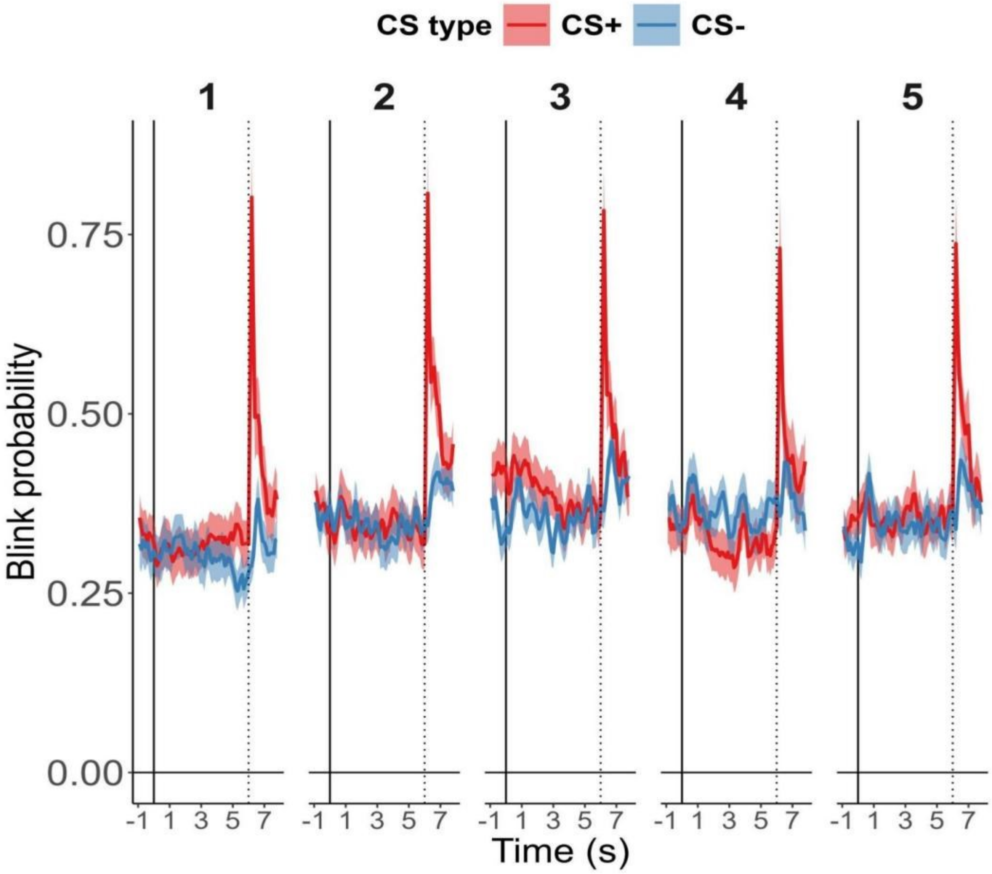


Figure 28. Time course of blink probability in response to CS+ and CS- across five blocks in the airpuff group(n = 25). Blink probability was analyzed in the 5.5-5.9 s time window following CS onset. The dotted vertical line at 6 s indicates the onset of the unconditioned stimulus (US). Shaded areas indicate ±1 standard error (SE).

3.3.6 Contingency awareness

Fear, arousal, and valence ratings. To examine the influence of contingency awareness on conditioned affective responses, a series of 2 (CS type: CS+, CS-) × 2 (Awareness Group: Aware, Unaware) mixed-design ANOVAs were conducted on fear, arousal, and valence ratings within the image group using data from Block 1.

For fear ratings, there was a significant main effect of CS type, $F(1, 50) = 11.09$, $p = 0.002$, $\eta^2_p = 0.18$, and a significant CS type × Awareness Group interaction, $F(1, 50) = 10.35$, $p = 0.002$, $\eta^2_p = 0.17$. The main effect of Awareness Group was not significant, $F(1, 50) = 0.16$, $p = 0.692$, $\eta^2_p < 0.01$. Follow-up comparisons indicated that the CS+ was rated as significantly more fearful than the CS- in the Aware group, $t(50) = 5.42$, $p < 0.001$, $d = 1.15$, but no significant difference was observed in the Unaware group, $t(50) = 0.07$, $p = 0.944$, $d = 0.01$ (see **Figure 29**).

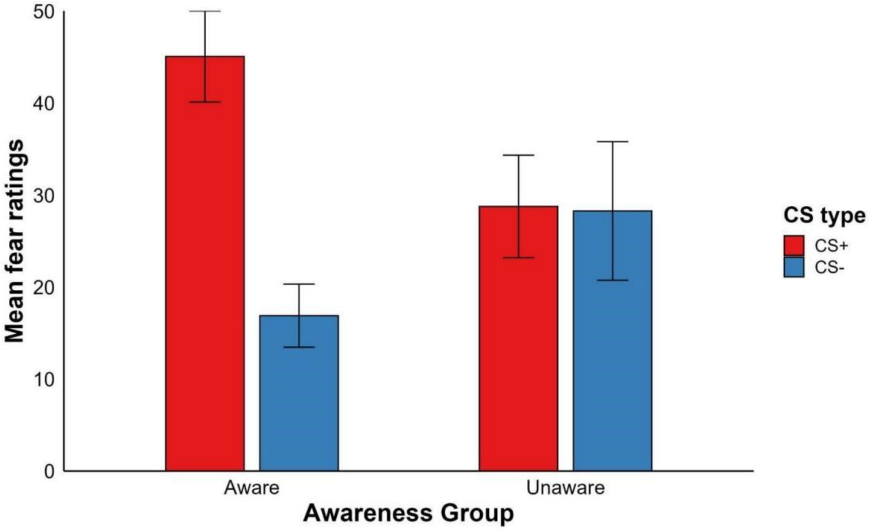


Figure 29. Mean fear ratings for CS+ and CS- across contingency awareness groups (aware: *n* = 33; unaware: *n* = 19) within the image US group (Block 1). Error bars represent ±1 standard error (SE).

For arousal ratings, the CS type × Awareness Group interaction was also significant, $F(1, 50) = 13.11$, $p < 0.001$, $\eta^2_p = 0.21$, whereas the main effects of CS type and Awareness Group were not significant (both *ps* > 0.05). In the Aware group, arousal ratings were significantly higher for CS+ than CS-, $t(50) = 4.33$, $p < 0.001$, $d = 0.82$. No significant difference was found in the Unaware group, $t(50) = -1.26$, $p = 0.213$, $d = -0.26$ (see **Figure 30**).

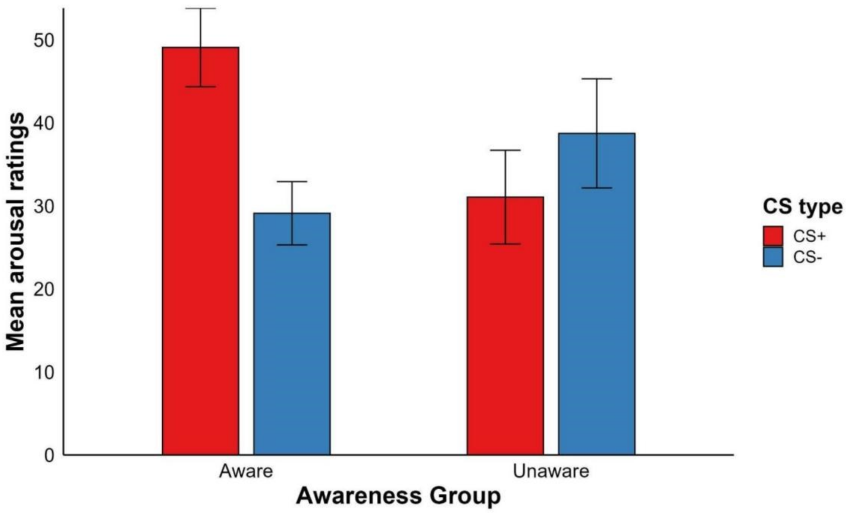


Figure 30. Mean arousal ratings for CS+ and CS- across contingency awareness groups (aware: *n* = 33; unaware: *n* = 19) within the image US group (Block 1).

For valence ratings, there was a significant main effect of CS type, *F*(1, 50) = 4.59, *p* = 0 .037, η^2_p= 0.08, and a significant CS type × Awareness Group interaction, *F*(1, 50) = 9.38, *p* = 0.004, η^2_p= 0.16. The main effect of Awareness Group was not significant, *F*(1, 50) = 0.18, *p* = 0.672, η^2_p< 0.01. Follow-up tests revealed that CS+ was rated as significantly more negative than CS- in the Aware group, *t*(50) = -4.30, *p* < 0.001, *d* = -0.74, but this difference was not observed in the Unaware group, *t*(50) = 0.58, *p* = 0.566, *d* = 0.14 (see **Figure 31**).

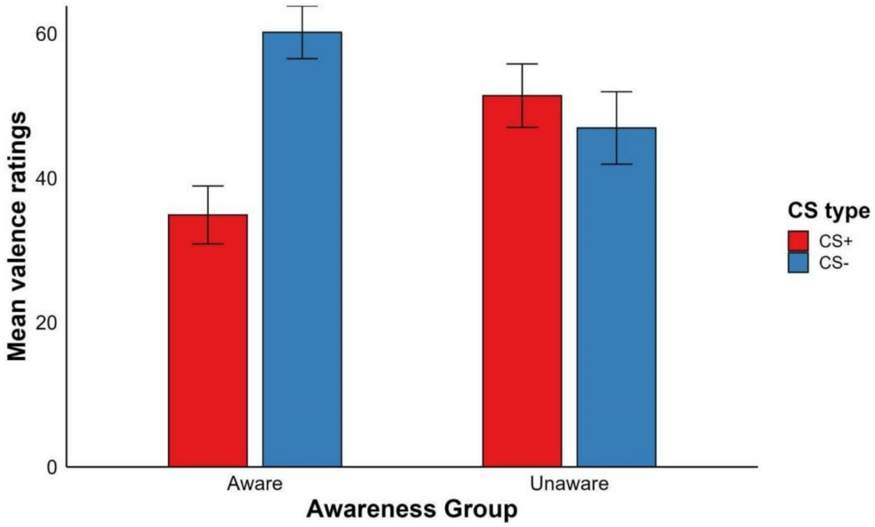


Figure 31. Mean valence ratings for CS+ and CS- across contingency awareness groups (aware: *n* = 33; unaware: *n* = 19) within the image US group (Block 1).

Skin conductance responses. The analysis of SCRs revealed no significant main effect of Awareness Group, $F(1, 50) = 0.33$, $p = 0.571$, $\eta^2_p < 0.01$, or CS type, $F(1, 50) = 1.51$, $p = 0.224$, $\eta^2_p = 0.03$. The interaction between Awareness Group and CS type was also not significant, $F(1, 50) = 0.27$, $p = 0.604$, $\eta^2_p < 0.01$.

Although the interaction was not significant, exploratory follow-up comparisons were conducted to examine within-group differences. In the Aware group, the difference in SCRs between CS+ and CS- was not statistically significant, $t(50) = 1.45$, $p = 0.153$, $d = 0.26$. Similarly, in the Unaware group, no significant difference was found, $t(50) = 0.44$, $p = 0.659$, $d = 0.10$ (see **Figure 32**).

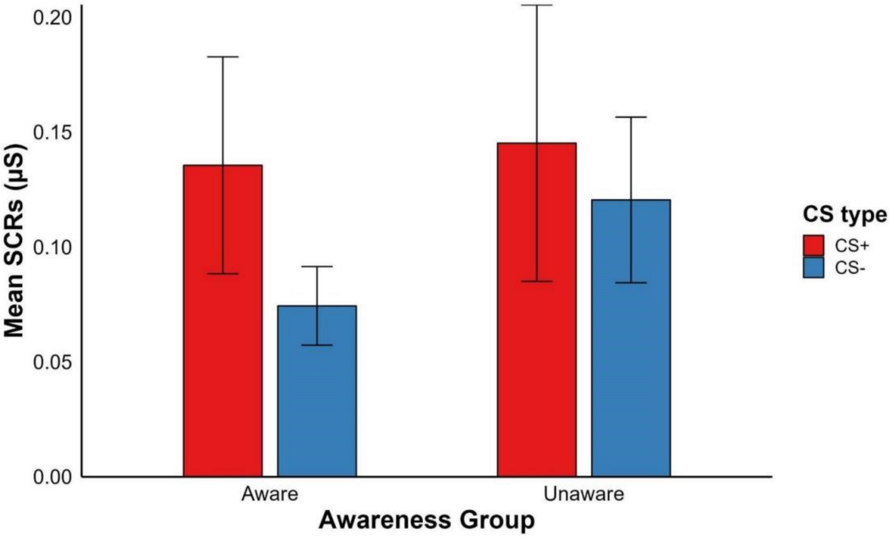


Figure 32. Mean SCRs for CS+ and CS- across contingency awareness groups (aware: $n = 33$; unaware: $n = 19$) within the image US group (Block 1).

Heart rate responses. The analysis of HR responses during the early phase (1.0-3.0 s post-CS onset) revealed no significant main effect of Awareness Group, $F(1, 49) = 0.17$, $p = 0.681$, $\eta^2_p < 0.01$, and no significant CS type × Awareness Group interaction, $F(1, 49) = 0.63$, $p = 0.430$, $\eta^2_p = 0.01$. The main effect of CS type approached significance, $F(1, 49) = 3.54$, $p = 0.066$, $\eta^2_p = 0.07$. Exploratory tests showed that in the Aware group, CS+ elicited a marked increase in heart rate, whereas CS- led to a slight decrease, resulting in a significant difference between the two conditions, $t(49) = 2.25$, $p = .029$, $d = 0.38$. In contrast, the unaware group showed no significant difference between CS+ and CS-, $t(49) = 0.67$, $p = .503$, $d = 0.17$ (see **Figure 33**).

The analysis of HR responses during the late window (3.0-5.5 s post-CS onset) revealed no significant main effect of Awareness Group, *F*(1, 49) = 0.04, *p* = 0.847, $\eta^2_p < 0.01$, and no significant CS type × Awareness Group interaction, *F*(1, 49) = 0.69, *p* = 0.409, $\eta^2_p = 0.01$. The main effect of CS type approached significance, *F*(1, 49) = 3.39, *p* = .072, $\eta^2_p = 0.06$. Exploratory post hoc tests revealed that, in the aware group during the late window, HR increased in response to CS+ but decreased in response to CS-, yielding a significant difference between conditions, *t*(49) = 2.25, *p* = 0.029, *d* = 0.38. No significant difference emerged in the Unaware group, *t*(49) = 0.63, *p* = 0.534, *d* = 0.16 (see **Figure 34**).

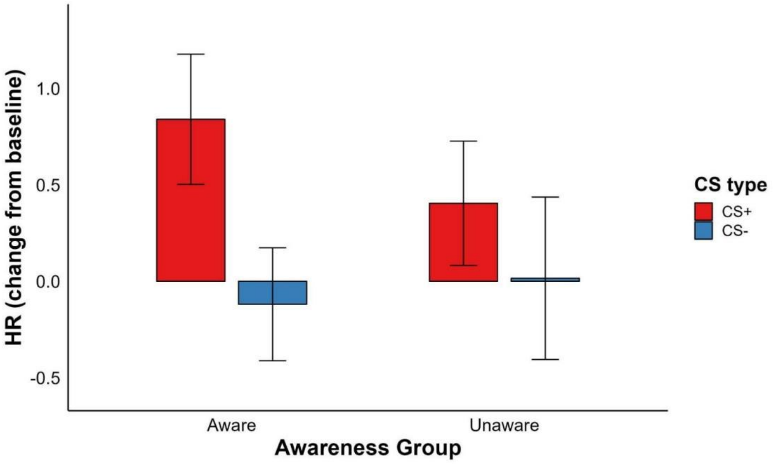


Figure 33. Mean HR changes from baseline during the acceleration phase (1.0-3.0 s post-CS onset) for CS+ and CS-, shown separately by contingency awareness groups (aware: *n* = 33; unaware: *n* = 19) within the image US group (Block 1).

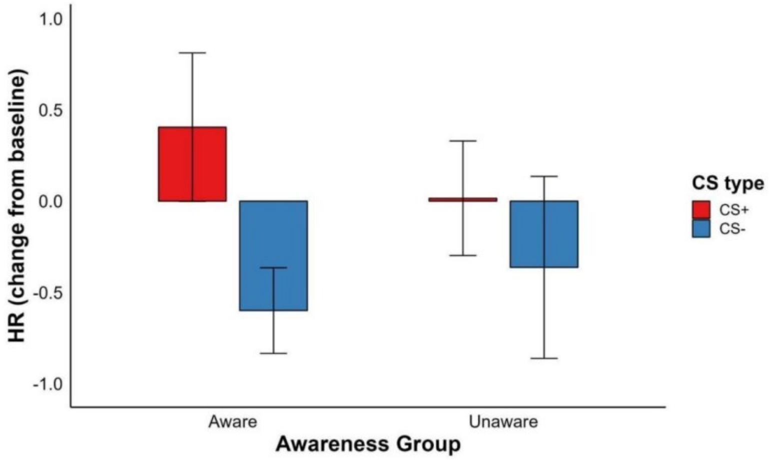


Figure 34. Mean HR changes from baseline during the deceleration phase (3.0-5.5 s post-CS onset) for CS+ and CS-, shown separately by contingency awareness groups (aware: *n* = 33; unaware: *n* = 19) within the image US group (Block 1).

3.3.7 Contingency awareness rate comparison between the three groups

The percentages of participants who acquired contingency awareness in the first block were 80.0% (28/35) in the airpuff US group, 63.5% (33/52) in the image US group, and 96.4% (27/28) in the shock US group. A Fisher's exact test revealed a significant difference in contingency awareness rates between the groups, $Z = 0.325$, $p = .002$, $V = 0.314$.

Pairwise post hoc comparisons showed that the image US group had a significantly lower awareness rate than the shock US group, $Z = 0.380$, $p < .001$, $V = 0.363$. The difference between the airpuff US group and the shock US group approached significance, $Z = 0.250$, $p = .066$, $V = 0.245$, whereas the difference between the airpuff US group and the image US group was not significant, $Z = -0.179$, $p = .151$, $V = 0.177$.

3.3.8 US valence and arousal ratings comparison between the three groups

Valence ratings of US differed significantly across modalities, as indicated by a Kruskal-Wallis test, $\chi^2(2) = 16.23$, $p < 0.001$, $\eta^2 = 0.13$, suggesting meaningful differences in perceived unpleasantness between aversive images, electric shocks, and airpuffs. Post hoc Dunn's tests (Holm-corrected) revealed that valence ratings were significantly lower for aversive images compared to both airpuffs ($Z = -3.90$, $p < 0.001$, $r = 0.37$) and electric shocks ($Z = -2.42$, $p = 0.031$, $r = 0.23$). No significant difference was found between electric shocks and airpuffs ($Z = 1.15$, $p = 0.249$, $r = 0.11$). **Table 9** presents the descriptive statistics and nonparametric test results for US affective ratings.

In contrast, arousal ratings did not significantly differ across US modalities, $\chi^2(2) = 4.54$, $p = 0.103$, $\eta^2 = 0.02$. These results suggest that while the perceived unpleasantness of US types varied, their overall arousing quality was rated similarly across participants.

Considering that US affective ratings were based on an unequal number of presentations across modalities (10 aversive images in the image US group vs. one airpuff and one electric shock in the other groups), we re-analyzed the data using only the first aversive image rating for the image US group. This re-analysis still revealed significant differences in valence ratings across modalities, $\chi^2(2) = 12.07$, $p = .002$, $\eta^2 = .09$. Detailed descriptive statistics and nonparametric test results are provided in **Appendix 6.9**. Post hoc Dunn's tests (Holm-corrected) showed that aversive images were rated as significantly more unpleasant than airpuffs (*Z* = 3.35, *p* = .002, *r* = .32), while the difference between aversive images and electric shocks approached significance (*Z* = -2.12, *p* = .067, *r* = .20). No significant difference was observed between airpuffs and electric shocks (*Z* = 0.95, *p* = .344, *r* = .09). In contrast, arousal ratings did not significantly differ across modalities, $\chi^2(2) = 5.68$, $p = .058$, $\eta^2 = .03$.

Table 9. Differences in valence and arousal ratings across US modalities (images, shocks, airpuffs).

	Images (*n* = 51)	Shocks (*n* = 28)	Airpuffs (*n* = 34)	Kruskal-Wallis χ^2 (2)	*p*	η^2
Valence	21.50 ± 12.18	32.99 ± 22.85	40.36 ± 23.76	16.23	**< 0.001***	0.13
Arousal	56.17 ± 26.65	47.54 ± 31.38	42.29 ± 30.75	4.54	0.103	0.02

Note. Values represent mean ± standard deviation. **p* < .05.

3.4 Discussion

This study examined how unconditioned stimulus (US) modality influences conditioned responses (CRs) across subjective, autonomic, and neural domains in human aversive conditioning. As hypothesized, electric shocks evoked the most robust and consistent CRs, reflected in significantly higher fear and arousal ratings and lower valence ratings for CS+ compared to CS-. These subjective effects were accompanied by increased skin conductance responses (SCRs), greater heart rate (HR) deceleration, greater stimulus-preceding negativity (SPN) negativity, enhanced late positive potential (LPP) amplitudes, and stronger alpha-band power suppression in response to CS+ compared to CS-. Airpuffs produced reliable HR deceleration and enhanced SPN for CS+, but elicited weaker subjective differentiation between CS+ and CS- compared to shocks and images, and did not elicit SCR discrimination. In contrast, aversive images yielded strong subjective differentiation between CS+ and CS-, yet failed to produce evident physiological or neural CRs.

Notably, within the image group, only participants who were aware of the CS-US contingency reported significantly greater fear, arousal, and negative valence ratings for CS+ compared to CS-, whereas unaware participants showed no such effects. It remains unclear whether physiological indicators of learning depend on contingency awareness, because SCR and HR conditioned responses were weak and non-significant even among aware participants.

3.4.1 US modality shapes strength and stability of evaluative conditioned responses

The results showed that the strength and stability of evaluative conditioned responses varied substantially depending on the modality of the US. Electric shocks elicited the most robust and stable responses, reflected in large and persistent differences in fear, arousal, and valence ratings between the CS+ and the CS- across the acquisition phase. Aversive images also produced strong subjective differentiation, with CS+ versus CS- differences broadly comparable to those observed in the shock group, albeit slightly weaker and more variable over time. In contrast, airpuffs elicited overall weaker and less consistent evaluative responses, with CS+-CS- differentiation declining across blocks, particularly in valence ratings. These findings suggest that both electric shocks and aversive images effectively supported the formation of evaluative conditioned responses, whereas airpuffs were considerably less effective.

At the end of the experiment, participants were re-exposed to the US used in their assigned group (aversive images, electric shocks, or airpuffs) and asked to rate the arousal and valence of the US. Notably, participants rated aversive images as more unpleasant than either shocks or airpuffs. Airpuffs were rated as the least unpleasant, which may account for their relatively weak evaluative effects. For fear ratings, the CS+ vs. CS- contrast yielded Cohen’s d = 0.64 for airpuffs, compared with 1.18 for images and 1.13 for shocks. Rapid habituation to airpuffs across repeated exposures may also have contributed to the decline in conditioned effects across acquisition blocks. A similar, although weaker, habituation pattern was observed in the US image group, whereas the CS+ vs. CS- difference remained fairly stable in the US shock group. Thus, shocks continued to be experienced as equally unpleasant at the end of the 55 trials, whereas conditioned responses in the other two modalities weakened over time with repeated US exposure. Post-experiment ratings indicated that images remained highly unpleasant overall and therefore produced evaluative responses comparable to shocks. Anticipation of a strongly unpleasant US increased fear, negative valence, and arousal ratings for the CS+. Because airpuffs were less aversive and habituated quickly, these evaluative effects were smaller. Taken together, these findings suggest that airpuffs may be less suitable for conditioning paradigms that require strong and enduring evaluative conditioning.

3.4.2 Divergent autonomic patterns reflect modality-specific processes

The autonomic CRs were strongly influenced by the US modality, with distinct patterns emerging across sympathetic and parasympathetic branches. In the shock group, CS+ consistently evoked greater SCRs than CS-, reflecting heightened sympathetic activation (Critchley, 2002). In contrast, no significant SCRs differentiation was observed in either the image or airpuff groups, suggesting that these stimuli did not reliably activate the sympathetic system. These findings are consistent with previous research showing that electric shocks produce greater SCR differentiation between CS+ and CS- than other aversive stimuli, such as trauma-related films, when used as the US (Ney, Schenker, et al., 2022). The robust SCR differentiation in the shock group likely reflects the direct nociceptive input and the high salience of painful stimulation, which strongly engages the sympathetic system and elicits heightened arousal (Kyle & McNeil, 2014). These findings further underscore the critical role of US modality in shaping sympathetic nervous system involvement in aversive conditioning.

HR responses also varied as a function of US modality. Both the shock and airpuff groups showed reliable bradycardia to CS+, indicating parasympathetic engagement (Battaglia et al., 2022). In the airpuff group, this deceleration emerged in both the early (1.0-3.0 s) and late (3.0-5.5 s) phases after CS onset and remained consistent across blocks, suggesting a sustained orienting or freezing response (Battaglia et al., 2024; Roelofs, 2017). One plausible explanation for this parasympathetic pattern lies in the type of emotional reaction elicited by airpuffs. As stimuli that intrude upon facial space, airpuffs typically provoke displeasure and startle-like arousal (Kauvar et al., 2025), a response profile more consistent with low-level disgust than with fear. Disgust, in turn, has been closely associated with parasympathetic activation, as reflected in cardiac deceleration (Klucken et al., 2012), which may explain the robust conditioned bradycardic responses observed in the airpuff group.

In the shock group, CS+ also elicited conditioned bradycardia, primarily during the late phase (3.0-5.5 s after CS onset), although the effect was less consistent across blocks. This may reflect concurrent sympathetic activation, as indicated by elevated SCRs, which could interfere with parasympathetic-driven HR slowing. In contrast, the image group showed minimal HR differentiation between CS+ and CS-. A modest acceleration to CS+ during the early phase (1.0-3.0 s after CS onset) may reflect an initial arousal or fear response, but the absence of late-phase deceleration suggests a lack of sustained parasympathetic involvement. This may be due to the emotional heterogeneity of aversive images, which can evoke diverse reactions such as fear, disgust, or sadness across individuals (Marchewka et al., 2014), thereby reducing the likelihood of consistent conditioned HR responses.

These results reveal modality-specific autonomic profiles, wherein different US modalities preferentially engage distinct branches of the autonomic nervous system: shocks predominantly activate the sympathetic system, airpuffs primarily recruit parasympathetic activity, and images elicit only early-phase sympathetic-related cardiac responses. These findings underscore the critical role of US modality in shaping autonomic dynamics during aversive learning.

3.4.3 EEG reveals modality-specific neural anticipation and evaluation

The EEG findings demonstrated that US modality significantly modulates conditioned neural responses. Across event-related potentials (ERPs) and time-frequency analyses, the most robust differentiation between CS+ and CS- was obtained in the group with electric shock US, followed by airpuffs. In contrast, no reliable EEG differentiation between CS+ and CS- was found in the image group.

The SPN is widely regarded as an electrophysiological marker of anticipatory attention and expectancy toward upcoming salient events (Baas et al., 2002; Böcker et al., 2001; Ferreira de Sá et al., 2019). In the present study, the SPN, recorded at Cz, showed significantly more negative amplitudes for CS+ relative to CS- in both the shock and airpuff groups. While the emergence of an SPN differential in the shock group is expected, given that shocks strongly activate the nociceptive system and reliably elicit fear-related defensive responses (Wiech & Tracey, 2013), the presence of a similar effect in the airpuff group requires further consideration. Although airpuffs are less aversive overall, they involve direct tactile stimulation to the body, often to the facial region, which is evolutionarily associated with immediate threat detection and defense. In addition, airpuffs tend to evoke relatively consistent affective reactions such as mild disgust across participants. These properties may be sufficient to engage anticipatory cortical mechanisms, as reflected in the observed SPN differential. In contrast, aversive images lack direct bodily impact and can elicit heterogeneous emotional reactions such as fear, disgust, or sadness (Marchewka et al., 2014), thereby reducing the likelihood of consistent anticipatory neural responses.

Beyond anticipatory processes indexed by the SPN, we also examined the LPP, which is widely regarded as an electrophysiological marker of sustained attentional and evaluative processing modulated by emotional arousal (Cuthbert et al., 2000; Hajcak et al., 2010; Schupp et al., 2000). For example, numerous studies have demonstrated that LPP amplitudes are larger for both pleasant and unpleasant stimuli compared to neutral stimuli (Pastor et al., 2008; Sabatinelli et al., 2007, 2013). Extending this evidence to aversive conditioning, studies have further shown that threat-predictive cues (CS+) elicit greater LPP amplitudes than safety cues (CS-) (Bruchmann et al., 2021; Paiva et al., 2020; Schindler et al., 2022; Shner-Livne et al., 2024). In the present study, a reliable CS+ > CS- LPP effect at POz was observed only in the shock group, consistent with prior findings (Diggs et al., 2022; Dou et al., 2023; Ferreira de Sá et al., 2019; Pastor et al., 2015; Pavlov & Kotchoubey, 2019; Sperl et al., 2021; Wiemer et al., 2021; Zhang et al., 2025), whereas no such differentiation was observed in the aversive image or eye-directed airpuff groups. This pattern likely reflects the high biological salience of electric shocks, which reliably provide sufficient arousal to support sustained evaluative processing. By contrast, aversive images and eye-directed airpuffs appear to lack the necessary salience and intensity to elicit comparable LPP differentiation. This interpretation is further supported by the finding that only the shock group exhibited clear conditioned sympathetic responses. Specifically, differential SCRs were observed between CS+ and CS- in the shock group, whereas no such effects emerged in the image or airpuff groups.

Although previous work has consistently shown that the LPP is larger for emotional than neutral pictures (Pastor et al., 2008; Sabatinelli et al., 2007, 2013), evidence that aversive images as US can reliably induce conditioned LPP effects is scarce, as available studies suffer from notable design and analytic limitations. For instance, one investigation employed an aversive conditioning paradigm with unpleasant images as US and examined LPP responses (Grant et al., 2015); however, it focused on the LPP elicited directly by the US themselves, rather than by the conditioned cues, thereby limiting its implications for conditioned LPP effects. Another study adopted a block-wise presentation of different CS categories, which were paired with neutral, unpleasant, or phobia-relevant pictures, and reported that cues signaling spider or unpleasant pictures elicited larger LPP amplitudes than safety cues (Michalowski et al., 2015). Yet, this block-wise CS presentation deviates substantially from general conditioning procedures, restricting the generalizability of its findings. With respect to airpuff stimulation, no study to date has demonstrated that an airpuff directed to the eye can serve as a US to elicit conditioned LPP responses. Although prior research has shown that an airblast to the larynx can successfully induce conditioned LPP effects (Seligowski et al., 2018), it is important to note that laryngeal airblasts are more intense, target a different anatomical site, and simulate a symptom-like choking sensation, which fundamentally distinguishes them from eye-directed airpuffs. Thus, current evidence provides only limited support for the capacity of aversive images and airpuffs to function as effective US in eliciting conditioned LPP responses.

The time-frequency analysis revealed that in the shock group, CS+ elicited significantly greater suppression of alpha-band power (9-15 Hz) at the C4 electrode compared to CS-. This effect likely reflects a conditioned response in the right somatosensory cortex, contralateral to the left wrist where the electric shocks were delivered. In contrast, no corresponding effect was observed at C3, the homologous site of C4 in the left hemisphere, indicating that the effect was lateralized to the contralateral hemisphere. This pattern is consistent with a previous study from our laboratory (Pavlov & Kotchoubey, 2019), in which electric shocks to the left wrist likewise elicited stronger suppression of lower beta activity (13-19 Hz) for CS+ compared to CS- in the right somatosensory region, while no such effect was found in the left region. Although alpha and beta suppression differ in their functional associations, both findings suggest selective engagement of the right-hemisphere somatosensory regions in threat prediction.

Importantly, no significant alpha suppression was observed at POz across any of the three US modalities. Alpha power suppression over parieto-occipital regions is commonly regarded as a neural marker of sustained attentional engagement to conditioned threat stimuli (Bacigalupo & Luck, 2022; Yin et al., 2020). In studies using aversive sounds as the US, this effect has proven particularly robust (Bacigalupo & Luck, 2022; Farkas et al., 2024; Friedl & Keil, 2020, 2021; Panitz et al., 2019; Yin et al., 2020), reliably eliciting alpha suppression across different CS modalities, including both visual and auditory CS (Farkas et al., 2024). Although electric shocks are typically regarded as highly salient, findings on parieto-occipital alpha suppression with shocks as the US have been less consistent. It has been reported that parieto-occipital alpha suppression distinguishing CS+ from CS- was not significant when shocks were inevitable, consistent with the results of the present shock group, but became significant when shocks were avoidable (Stegmann et al., 2024). In the avoidable case, CS+ carries higher motivational salience, as participants must prepare defensive responses, thereby enhancing attention and facilitating alpha suppression. By contrast, significant alpha suppression was observed even under inevitable shocks, but this effect occurred in a study that required participants to provide US expectancy ratings during CS presentation (Dou et al., 2023). This task likely amplified attentional orienting and working memory load and thus inflated alpha suppression. Taken together, these findings suggest that alpha suppression effects with shocks as the US remain inconclusive and may depend strongly on threat avoidability and task demands. For the airpuff and image groups, the absence of alpha suppression is most likely due to their relatively low aversive intensity.

Overall, the EEG results indicate a hierarchy in conditioned neural responses: anticipatory mechanisms reflected in the SPN can be recruited even by relatively mild aversive stimuli such as airpuffs, whereas more sustained attentional and evaluative processes indexed by the LPP and alpha suppression emerge primarily when the US carries high biological salience, such as electric shocks. Even in this case, however, parieto-occipital alpha suppression requires further investigation. This underscores that the nature of the US critically shapes both the presence and the strength of conditioned neural responses.

The modality-specific effects observed across subjective, autonomic, and neural measures likely result from the combined influence of several factors, including biological salience, sensory modality, emotional quality such as fear versus disgust, and the physical intensity of the US. Future research should aim to clarify how each of these factors contributes, either independently or in interaction with others, to the development of CRs across subjective experience, autonomic function, and neural processing.

3.4.4 Contingency awareness shapes subjective responses while physiological effects remain inconclusive

The current findings demonstrate that contingency awareness plays a critical role in the formation of evaluative conditioned responses. Participants who correctly identified the CS-US association rated the CS+ as more fearful, more arousing, and more negatively valenced than the CS-, whereas participants who failed to acquire this association showed no such differentiation. This pattern is highly consistent with the results of Study 1 (Li et al., 2025).

However, the relationship between contingency awareness and physiological indices of aversive conditioning remains unclear. For SCRs, no significant interaction was found between the awareness group and CS type, indicating comparable SCRs patterns across groups. Although the aware group exhibited numerically greater differences between CS+ and CS- compared to the unaware group, exploratory post hoc analyses revealed that neither group showed significant SCR differentiation between the two stimuli.

For HR analyses of both the early and late phases similarly yielded no significant interaction between group and CS type, suggesting that HR responses were generally comparable across awareness groups. However, exploratory analyses suggested that in the aware group, HR acceleration was significantly greater for CS+ than for CS- in both phases. In contrast, the unaware group showed no significant HR differences between CS types during either phase. These exploratory results suggest that contingency awareness may influence HR dynamics, although the relationship remains inconclusive and warrants further investigation.

3.4.5 Limitation

A key limitation of this study is that the relationship between contingency awareness and aversive conditioning was assessed exclusively in the image group. As a result, it remains unclear whether similar patterns would emerge with other US modalities, such as electric shocks or airpuffs - the hypothesis that could not been tested in the current study due to very high prevalence of participants who became aware in the US shock and airpuff groups making the comparison between contingency aware and unaware participants impossible. Future research should examine whether the influence of contingency awareness extends across different US types.

3.5 Conclusion

This study demonstrates that CRs are strongly shaped by the modality of the US, with different US types engaging distinct affective, autonomic, and neural systems. Electric shocks elicited the most robust and consistent CRs across all domains, including heightened sympathetic activity and pronounced neural anticipation and evaluation. Aversive images supported strong subjective differentiation but failed to elicit consistent physiological or neural responses. Airpuffs primarily engaged parasympathetic activity and anticipatory neural responses, though their subjective evaluation was weaker and less stable.

Contingency awareness emerged as a key factor influencing subjective affective learning, though its role in physiological responses remains inconclusive. These findings contribute to a more nuanced understanding of aversive conditioning and offer practical guidance for selecting US in experimental designs. For studies aiming to elicit strong sympathetic and cortical responses, electric shocks appear most effective. When the focus is on affective discrimination, both shocks and aversive images are appropriate. In contrast, airpuffs may be better suited for paradigms emphasizing parasympathetic or orienting responses, particularly those involving heart rate measures.

4 General Discussion

This dissertation investigated two central questions in the field of aversive conditioning: whether contingency awareness is necessary for the acquisition of conditioned fear, and how different unconditioned stimulus (US) modalities shape conditioned responses (CRs). To address these questions, two complementary studies were conducted.

The findings converge on two key insights. First, contingency awareness proved essential for the development of subjective fear responses, as only participants who recognized the CS-US relationship exhibited reliable affective differentiation between CS+ and CS-. However, its role in physiological and neural domains was less conclusive. Second, the stimulus modality of CS and US substantially shaped both the acquisition of contingency awareness and the pattern of CRs. Contingency awareness was less likely to develop with auditory CS, which consisted of 1,000 Hz and 1,400 Hz tones, than with visual CS, which consisted of circle and diamond shapes. In contrast, auditory US (loud noise) and nociceptive US (electric shocks) facilitated contingency awareness compared to visual US (aversive images). Moreover, the modality of the US shaped both the overall strength of aversive learning and the specific response systems engaged: electric shocks elicited the most robust CRs across subjective, autonomic, and neural systems; aversive images primarily yielded evaluative CRs; and airpuffs produced weaker evaluative CRs, accompanied by reliable heart rate deceleration and anticipatory cortical activity.

4.1 Role of Contingency Awareness in Fear Conditioning

The findings of this research offer converging support for the claim that contingency awareness is a necessary condition for the acquisition of subjective fear responses. In both studies, only participants who were aware of the CS-US contingency exhibited reliable affective differentiation between the CS+ and the CS-. This pattern consistently emerged across both online and in-lab experiments and across diverse participant samples. These results align with the conclusions of Lovibond and Shanks, who proposed that Pavlovian conditioning relies on a propositional learning mechanism that enables the formation of conscious expectations about stimulus contingencies, thereby supporting CRs (Lovibond & Shanks, 2002). Similarly, research on evaluative conditioning has demonstrated that the co-occurrence of a CS and an emotional outcome typically alters the affective value of the CS only when participants are aware of the contingency (De Houwer et al., 2001; Moran et al., 2023). Further support comes from a meta-analysis by Mertens and Engelhard (Mertens & Engelhard, 2020), which found no convincing evidence for fear conditioning without contingency awareness. Across 41 studies, methodological flaws were common, and higher-quality studies were less likely to support unaware fear learning.

By contrast, in the present research, the relationship between contingency awareness and physiological measures of fear conditioning, including SCRs and HR, remains inconclusive. In Study 2, contingency awareness did not significantly influence physiological indices of fear learning. Although exploratory analyses indicated that HR acceleration and deceleration differed between CS+ and CS- within the aware group, the interaction between CS type and awareness status was not statistically significant. Similarly, for SCRs, no reliable differentiation between CS+ and CS- was observed in either awareness group, and no significant interaction with awareness status was found. Importantly, these results do not imply that physiological responses in fear conditioning can occur independently of contingency awareness. Neither the aware nor the unaware group exhibited strong or consistent physiological effects within the aversive image group, suggesting that the conditioning procedure may not have produced robust physiological learning overall. Although previous studies have reported conditioned physiological responses, such as fear-potentiated startle and SCRs, without explicit awareness (Jovanovic et al., 2006; Raio et al., 2012; Schultz & Helmstetter, 2010), these findings remain contentious due to methodological limitations (Lovibond & Shanks, 2002; Mertens & Engelhard, 2020). Taken together, the relationship between contingency awareness and conditioned physiological responses remains uncertain.

In addition, prior research has suggested that certain neural markers of fear learning, such as amygdala activation (Knight et al., 2009; Tabbert et al., 2011) and specific EEG components (Steinberg et al., 2012; Wong et al., 2004), may emerge even in the absence of full conscious awareness of the CS-US contingency. Therefore, the relationship between contingency awareness and conditioned neural responses, such as the SPN, LPP, and alpha suppression, warrants further investigation.

In summary, within the domain of aversive conditioning, contingency awareness appears to be a critical prerequisite for evaluative conditioned responses. However, whether contingency awareness constitutes a necessary condition for conditioned physiological or neural responses remains an open question.

4.2 Role of Stimulus Modality in Aversive Conditioning

The present findings demonstrate that stimulus modality critically influences both the acquisition of contingency awareness and the pattern of CRs. In Study 1, participants were less likely to acquire contingency awareness when auditory CS were used compared to visual CS (57.3% vs. 75.3%). This difference may reflect the greater difficulty of discriminating auditory stimuli relative to visual ones, which hinders the detection of CS-US contingencies. Alternatively, auditory CS may have distracted participants from the visual display, thereby reducing opportunities to encode the predictive relationship.

By contrast, the auditory US demonstrated the reverse pattern. Participants in the loud-noise US group acquired contingency awareness at a significantly higher rate than those in the image US group (94.7% vs. 75.3%). Loud noise elicited stronger negative emotional reactions than aversive images and was difficult to ignore without violating task requirements, which likely compelled participants to attend to the auditory US and recognize its predictive signal (CS+). In Study 2, US modality again significantly influenced awareness. The awareness rate was 63.5% in the image US group, 80.0% in the airpuff US group, and 96.4% in the shock US group. The awareness rate was significantly lower in the image group than in the shock group, and the difference between the airpuff group and the shock group approached significance (p = .066), whereas the image group and the airpuff group did not differ. The particularly high rate of awareness in the shock US group can be attributed to the strong biological salience of electric shocks, which directed participants' attention toward the US and thereby facilitated clearer awareness of the CS-US association. Taken together, these results indicate that both CS and US modality substantially shape the acquisition of contingency awareness.

Beyond awareness, US modality also influenced evaluative, physiological, and neural CRs. Across both studies, loud noise, electric shocks, and aversive images reliably led to strong evaluative CRs, as reflected in affective ratings. However, Study 2 revealed a crucial dissociation: although both shocks and aversive images produced strong subjective CRs, only shocks were associated with robust conditioned physiological and neural effects. Specifically, the shock group exhibited reliable differential SCRs, HR deceleration, more negative SPN amplitudes, larger LPP amplitudes, and stronger alpha-band suppression to CS+, whereas the image group failed to produce consistent physiological or neural differentiation. The airpuff group, in turn, displayed a distinctive pattern. Compared to the shock and image groups, the airpuff group showed weaker evaluative CRs, yet reliably elicited HR deceleration and a pronounced conditioned SPN to CS+, suggesting anticipatory cortical engagement. However, the airpuff group did not exhibit conditioned SCRs or LPP effects.

In summary, the combined results of Studies 1 and 2 highlight that stimulus modality is a key factor shaping aversive conditioning, influencing not only the likelihood of acquiring contingency awareness but also the strength and profile of CRs across subjective, physiological, and neural domains.

4.3 Future Directions

First, future research should employ more sensitive and multidimensional measures of contingency awareness. In some cases, participants may acquire only partial knowledge of CS-US contingencies, suggesting that a strict binary classification into "aware" and "unaware" groups may be overly simplistic. A more refined approach would be to conceptualize awareness as a continuum, for example by incorporating confidence ratings or other graded indices, thereby providing a more nuanced characterization of individual differences in awareness and their relationship to conditioned responses.

Second, when aversive images are used as the US, physiological responses tend to be relatively weak, resulting in non-significant effects in both the aware and unaware groups, which may reflect a floor effect. In addition, aversive images generally fail to elicit robust neural responses. To better investigate the influence of contingency awareness on aversive conditioning, a more suitable strategy would be to employ US with higher biological salience, such as loud noise or electric shocks. Notably, loud noise has been shown to elicit clear conditioned physiological and neural indices in previous studies employing EEG measures (Bacigalupo & Luck, 2022; Sperl et al., 2016) and may therefore represent an ideal choice for future research.

4.4 Conclusion

This dissertation demonstrates that contingency awareness is essential for the acquisition of subjective fear responses, although its influence on physiological and neural outcomes remains less clear. The development of awareness was shaped by several procedural factors, including instruction type, the use of online ratings, engagement in concurrent tasks, and the modality of both the CS and the US. These factors primarily influenced awareness acquisition by modulating attention to the CS-US association.

The findings also underscore the critical role of US modality in shaping conditioned responses. Electric shocks produced the most robust effects across subjective, autonomic, and neural systems. In contrast, aversive images primarily gave rise to subjective responses, whereas airpuffs yielded weaker evaluative CRs but reliably evoked heart rate deceleration and anticipatory cortical activity.

By jointly elucidating the pivotal role of contingency awareness and the influence of stimulus modality, this dissertation advances a more nuanced understanding of the mechanisms underlying human aversive conditioning and provides important insights for clinical interventions targeting anxiety and fear-related disorders.

5 References

Abramson, L., Callaghan, B. L., Silvers, J. A., Choy, T., VanTieghem, M., Vannucci, A., Fields, A., & Tottenham, N. (2024). The effects of parental presence on amygdala and mPFC activation during fear conditioning: An exploratory study. *Developmental Science*, *27*(6), e13505. https://doi.org/10.1111/desc.13505

Acheson, D. T., Forsyth, J. P., Prenoveau, J. M., & Bouton, M. E. (2007). Interoceptive fear conditioning as a learning model of panic disorder: An experimental evaluation using 20% CO(2)-enriched air in a non-clinical sample. *Behaviour Research and Therapy*, *45*(10), 2280–2294. https://doi.org/10.1016/j.brat.2007.04.008

Allen, M. T., Handy, J. D., Blankenship, M. R., & Servatius, R. J. (2018). The distressed (Type D) personality factor of social inhibition, but not negative affectivity, enhances eyeblink conditioning. *Behavioural Brain Research*, *345*, 93–103. https://doi.org/10.1016/j.bbr.2018.02.035

Baas, J. M. P., Kenemans, J. L., Böcker, K. B. E., & Verbaten, M. N. (2002). Threat-induced cortical processing and startle potentiation. *Neuroreport*, *13*(1), 133–137. https://doi.org/10.1097/00001756-200201210-00031

Babel, P., Adamczyk, W., Swider, K., Bajcar, E. A., Kicman, P., & Lisinska, N. (2018). How classical conditioning shapes placebo analgesia: Hidden versus open conditioning. *Pain Medicine, 19*(6), 1156–1169. https://doi.org/10.1093/pm/pnx177

Bach, D. R., Sporrer, J., Abend, R., Beckers, T., Dunsmoor, J. E., Fullana, M. A., Gamer, M., Gee, D. G., Hamm, A., Hartley, C. A., Herringa, R. J., Jovanovic, T., Kalisch, R., Knight, D. C., Lissek, S., Lonsdorf, T. B., Merz, C. J., Milad, M., Morriss, J., … Schiller, D. (2023). Consensus design of a calibration experiment for human fear conditioning. *Neuroscience and Biobehavioral Reviews*, *148*, 105146. https://doi.org/10.1016/j.neubiorev.2023.105146

Bacigalupo, F., & Luck, S. J. (2018). Event-related potential components as measures of aversive conditioning in humans. *Psychophysiology*, *55*(4), e13015. https://doi.org/10.1111/psyp.13015

Bacigalupo, F., & Luck, S. J. (2022). Alpha-band EEG suppression as a neural marker of sustained attentional engagement to conditioned threat stimuli. *Social Cognitive and Affective Neuroscience*, *17*(12), 1101–1117. https://doi.org/10.1093/scan/nsac029

Baeuchl, C., Hoppstädter, M., Meyer, P., & Flor, H. (2019). Contingency awareness as a prerequisite for differential contextual fear conditioning. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(4), 811–828. https://doi.org/10.3758/s13415-018-00666-z

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1990). Flavor-flavor and color-flavor conditioning in humans. *Learning and Motivation*, *21*, 434–455.

https://doi.org/10.1016/0023-9690(90)90025-J

Barrett, S. T., McNealy, K. R., Knabel, M. L., Burrichter, R. M., Steck, K. A., & Bevins, R. A. (2025). The inside story: Interoceptive Pavlovian conditioning with the nicotine stimulus. *Journal of Experimental Psychology. Animal Learning and Cognition*, *51*(2), 61–72. https://doi.org/10.1037/xan0000393

Battaglia, S., Nazzi, C., Lonsdorf, T. B., & Thayer, J. F. (2024). Neuropsychobiology of fear-induced bradycardia in humans: Progress and pitfalls. *Molecular Psychiatry*, *29*(12), 3826–3840. https://doi.org/10.1038/s41380-024-02600-x

Battaglia, S., Orsolini, S., Borgomaneri, S., Barbieri, R., Diciotti, S., & di Pellegrino, G. (2022). Characterizing cardiac autonomic dynamics of fear learning in humans. *Psychophysiology*, *59*(12), e14122. https://doi.org/10.1111/psyp.14122

Berg, F., Margraf, J., & Wannemüller, A. (2023). Calibrating your own fears: Feasibility of a remote fear conditioning paradigm with semi-subjective stimulus calibration and differences in fear learning. *Learning & Behavior*, *51*(3), 246–261. https://doi.org/10.3758/s13420-022-00545-1

Biferno, M. A., & Dawson, M. E. (1977). The Onset of Contingency Awareness and Electrodermal Classical Conditioning: An Analysis of Temporal Relationships during Acquisition and Extinction. *Psychophysiology*, *14*(2), 164–171. https://doi.org/10.1111/j.1469-8986.1977.tb03370.x

Biggs, E. E., Timmers, I., Meulders, A., Vlaeyen, J. W. S., Goebel, R., & Kaas, A. L. (2020). The neural correlates of pain-related fear: A meta-analysis comparing fear conditioning studies using painful and non-painful stimuli. *Neuroscience and Biobehavioral Reviews*, *119*, 52–65. https://doi.org/10.1016/j.neubiorev.2020.09.016

Björkstrand, J., Pine, D. S., & Frick, A. (2022). Evaluating an internet-delivered fear conditioning and extinction protocol using response times and affective ratings. *Scientific Reports*, *12*(1), 4014. https://doi.org/10.1038/s41598-022-07999-3

Blumenthal, T. D. (2015). Presidential Address 2014: The more-or-less interrupting effects of the startle response. *Psychophysiology*, *52*(11), 1417–1431. https://doi.org/10.1111/psyp.12506

Böcker, K. B., Baas, J. M., Kenemans, J. L., & Verbaten, M. N. (2001). Stimulus-preceding negativity induced by fear: A manifestation of affective anticipation. *International Journal of Psychophysiology, 43*(1), 77–90. https://doi.org/10.1016/S0167-8760(01)00180-5

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*(2), 201–206. https://doi.org/10.1016/j.jbtep.2012.08.003

Bolaram, A., Coe, T. E., Power, J. M., & Cheng, D. T. (2020). Awareness and differential eyeblink conditioning: Effects of manipulating auditory CS frequencies. *Learning & Memory, 27*(2), 78–82. https://doi.org/10.1101/lm.050146.119

Bowman, H., Brooks, J. L., Hajilou, O., Zoumpoulaki, A., & Litvak, V. (2020). Breaking the circularity in circular analyses: Simulations and formal treatment of the flattened average approach. *PLoS Computational Biology*, *16*(11), e1008286. https://doi.org/10.1371/journal.pcbi.1008286

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Bruchmann, M., Schindler, S., Heinemann, J., Moeck, R., & Straube, T. (2021). Increased early and late neuronal responses to aversively conditioned faces across different attentional conditions. *Cortex, 142*, 332–341. https://doi.org/10.1016/j.cortex.2021.07.003

Burns, A., & Zaudig, M. (2002). Mild cognitive impairment in older people. *The Lancet*, *360*(9349), 1963–1965. https://doi.org/10.1016/S0140-6736(02)11920-9

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Ceunen, E., Zaman, J., Herssens, N., Van Oudenhove, L., Bogaerts, K., Ly, H. G., Coen, S. J., Tack, J., Vlaeyen, J. W. S., & Van Diest, I. (2015). Startle responding in the context of visceral pain. *International Journal of Psychophysiology, 98*(1), 128–134. https://doi.org/10.1016/j.ijpsycho.2015.07.009

Chan, C. K., & Lovibond, P. F. (1996). Expectancy bias in trait anxiety. *Journal of Abnormal Psychology*, *105*(4), 637–647. https://doi.org/10.1037//0021-843x.105.4.637

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science, 280*(5360), 77–81. https://doi.org/10.1126/science.280.5360.77

Constantinou, E., Purves, K. L., McGregor, T., Lester, K. J., Barry, T. J., Treanor, M., Craske, M. G., & Eley, T. C. (2021). Measuring fear: Association among different measures of fear learning. *Journal of Behavior Therapy and Experimental Psychiatry*, *70*, 101618. https://doi.org/10.1016/j.jbtep.2020.101618

Cosand, L. D., Cavanagh, T. M., Brown, A. A., Courtney, C. G., Rissling, A. J., Schell, A. M., & Dawson, M. E. (2008). Arousal, working memory, and conscious awareness in contingency learning. *Consciousness and Cognition*, *17*(4), 1105–1113. https://doi.org/10.1016/j.concog.2008.04.007

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *The Neuroscientist*, *8*(2), 132–142. https://doi.org/10.1177/107385840200800209

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111. https://doi.org/10.1016/S0301-0511(99)00044-7

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

Dawson, M. E., & Biferno, M. A. (1973). Concurrent measurement of awareness and electrodermal classical conditioning. *Journal of Experimental Psychology*, *101*(1), 55–62. https://doi.org/10.1037/h0035524

Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion*, *7*(4), 755–766. https://doi.org/10.1037/1528-3542.7.4.755

Dawson, M. E., Schell, A. M., & Banis, H. T. (1986). Greater Resistance to Extinction of Electrodermal Responses Conditioned to Potentially Phobic CSs: A Noncognitive Process? *Psychophysiology*, *23*(5), 552–561. https://doi.org/10.1111/j.1469-8986.1986.tb00673.x

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869. https://doi.org/10.1037/0033-2909.127.6.853

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Dieterich, M., & Brandt, T. (2024). Central vestibular networking for sensorimotor control, cognition, and emotion. *Current Opinion in Neurology*, *37*(1), 74–82. https://doi.org/10.1097/WCO.0000000000001233

Diggs, H. A., Rabinovich, N. E., & Gilbert, D. G. (2022). Facilitated extinction of conditioned fear responses by delta 9-tetrahyrdrocannabidol in humans: A pilot study. *Human Psychopharmacology*, *37*(6), e2853. https://doi.org/10.1002/hup.2853

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2007). Reinstatement of conditioned responses in human differential fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(3), 237–251. https://doi.org/10.1016/j.jbtep.2006.04.001

Dou, H., Lei, Y., Pan, Y., Li, H., & Astikainen, P. (2023). Impact of observational and direct learning on fear conditioning generalization in humans. *Progress in Neuro-Psychopharmacology &*

Biological Psychiatry, *121*, 110650. https://doi.org/10.1016/j.pnpbp.2022.110650

Eelen, P. (2018). Classical Conditioning: Classical Yet Modern. *Psychologica Belgica*, *58*(1), 196–211. https://doi.org/10.5334/pb.451

Farkas, A. H., Ward, R. T., Gilbert, F. E., Pouliot, J., Chiasson, P., McIlvanie, S., Traiser, C., Riels, K., Mears, R., & Keil, A. (2024). Auditory aversive generalization learning prompts threat-specific changes in alpha-band activity. *Cerebral Cortex, 34*(3), bhae099. https://doi.org/10.1093/cercor/bhae099

Felmingham, K. L., Caruana, J. M., Miller, L. N., Ney, L. J., Zuj, D. V., Hsu, C. M. K., Nicholson, E., To, A., & Bryant, R. A. (2021). Lower estradiol predicts increased reinstatement of fear in women. *Behaviour Research and Therapy*, *142*, 103875. https://doi.org/10.1016/j.brat.2021.103875

Ferreira de Sá, D. S., Michael, T., Wilhelm, F. H., & Peyk, P. (2019). Learning to see the threat: Temporal dynamics of ERPs of motivated attention in fear conditioning. *Social Cognitive and Affective Neuroscience*, *14*(2), 189–203. https://doi.org/10.1093/scan/nsy103

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences*, *17*(6), 791–802.

Friedl, W. M., & Keil, A. (2020). Effects of Experience on Spatial Frequency Tuning in the Visual System: Behavioral, Visuocortical, and Alpha-band Responses. *Journal of Cognitive Neuroscience*, *32*(6), 1153–1169. https://doi.org/10.1162/jocn_a_01524

Friedl, W. M., & Keil, A. (2021). Aversive Conditioning of Spatial Position Sharpens Neural Population-Level Tuning in Visual Cortex and Selectively Alters Alpha-Band Activity. *The Journal of Neuroscience*, *41*(26), 5723–5733. https://doi.org/10.1523/JNEUROSCI.2889-20.2021

Fullana, M. A., Harrison, B. J., Soriano-Mas, C., Vervliet, B., Cardoner, N., Àvila-Parcet, A., & Radua, J. (2016). Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies. *Molecular Psychiatry*, *21*(4), 500–508. https://doi.org/10.1038/mp.2015.88

Furuyama, T., Imayoshi, A., Iyobe, T., Ono, M., Ishikawa, T., Ozaki, N., Kato, N., & Yamamoto, R. (2023). Multiple factors contribute to flight behaviors during fear conditioning. *Scientific Reports*, *13*(1), 10402. https://doi.org/10.1038/s41598-023-37612-0

Geller, J., Winn, M. B., Mahr, T., & Mirman, D. (2020). GazeR: A Package for Processing Gaze Position and Pupil Size Data. *Behavior Research Methods*, *52*(5), 2232–2255. https://doi.org/10.3758/s13428-020-01374-8

Gershman, S. J., & Hartley, C. A. (2015). Individual differences in learning predict the return of fear. *Learning & Behavior*, *43*(3), 243–250. https://doi.org/10.3758/s13420-015-0176-z

Glenn, C. R., Klein, D. N., Lissek, S., Britton, J. C., Pine, D. S., & Hajcak, G. (2012). The development of fear learning and generalization in 8-13 year-olds. *Developmental Psychobiology*, *54*(7), 675–684. https://doi.org/10.1002/dev.20616

Glenn, C. R., Lieberman, L., & Hajcak, G. (2012). Comparing electric shock and a fearful screaming face as unconditioned stimuli for fear learning. *International Journal of Psychophysiology*, *86*(3), 214–219. https://doi.org/10.1016/j.ijpsycho.2012.09.006

Grant, D. M., Judah, M. R., White, E. J., & Mills, A. C. (2015). Worry and Discrimination of Threat and Safety Cues: An Event-Related Potential Investigation. *Behavior Therapy*, *46*(5), 652–660. https://doi.org/10.1016/j.beth.2014.09.015

Grillon, C., Baas, J. M. P., Cornwell, B., & Johnson, L. (2006). Context conditioning and behavioral avoidance in a virtual reality environment: Effect of predictability. *Biological Psychiatry*, *60*(7), 752–759. https://doi.org/10.1016/j.biopsych.2006.03.072

Grillon, C., Baas, J. P., Lissek, S., Smith, K., & Milstein, J. (2004). Anxious responses to predictable and unpredictable aversive events. *Behavioral Neuroscience*, *118*(5), 916–924. https://doi.org/10.1037/0735-7044.118.5.916

Gundlach, C., Müller, M. M., Nierhaus, T., Villringer, A., & Sehm, B. (2017). Modulation of Somatosensory Alpha Rhythm by Transcranial Alternating Current Stimulation at Mu-Frequency. *Frontiers in Human Neuroscience*, *11*, 432. https://doi.org/10.3389/fnhum.2017.00432

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Harris, J. A. (2025). Modelling the acquisition of Pavlovian conditioning. *Neurobiology of Learning and Memory*, *219*, 108059. https://doi.org/10.1016/j.nlm.2025.108059

Harvie, D. S., Meulders, A., Madden, V. J., Hillier, S. L., Peto, D. K., Brinkworth, R., & Moseley, G. L. (2016). When touch predicts pain: Predictive tactile cues modulate perceived intensity of painful stimulation independent of expectancy. *Scandinavian Journal of Pain*, *11*, 11–18. https://doi.org/10.1016/j.sjpain.2015.09.007

Hauck, A., Michael, T., & Ferreira de Sá, D. S. (2022). Fear learning and generalization during pandemic fear: How COVID-19-related anxiety affects classical fear conditioning with traumatic film clips. *Journal of Psychiatric Research*, *155*, 90–99. https://doi.org/10.1016/j.jpsychires.2022.07.068

Hennings, A. C., Cooper, S. E., Lewis-Peacock, J. A., & Dunsmoor, J. E. (2022). Pattern analysis of neuroimaging data reveals novel insights on threat learning and extinction in humans. *Neuroscience and Biobehavioral Reviews*, *142*, 104918. https://doi.org/10.1016/j.neubiorev.2022.104918

Icenhour, A., Kattoor, J., Benson, S., Boekstegers, A., Schlamann, M., Merz, C. J., Forsting, M., & Elsenbruch, S. (2015). Neural circuitry underlying effects of context on human pain-related fear extinction in a renewal paradigm. *Human Brain Mapping*, *36*(8), 3179–3193. https://doi.org/10.1002/hbm.22837

Jenewein, J., Moergeli, H., Sprott, H., Honegger, D., Brunner, L., Ettlin, D., Grillon, C., Bloch, K., Brügger, M., Schwegler, K., Schumacher, S., & Hasler, G. (2013). Fear-learning deficits in subjects with fibromyalgia syndrome? *European Journal of Pain, 17*(9), 1374–1384. https://doi.org/10.1002/j.1532-2149.2013.00300.x

Jordan, W. P., Todd, T. P., Bucci, D. J., & Leaton, R. N. (2015). Habituation, latent inhibition, and extinction. *Learning & Behavior*, *43*(2), 143–152. https://doi.org/10.3758/s13420-015-0168-z

Jovanovic, T., Norrholm, S. D., Keyes, M., Fiallos, A., Jovanovic, S., Myers, K. M., Davis, M., & Duncan, E. J. (2006). Contingency awareness and fear inhibition in a human fear-potentiated startle paradigm. *Behavioral Neuroscience*, *120*(5), 995–1004. https://doi.org/10.1037/0735-7044.120.5.995

Kampa, M., Stark, R., & Klucken, T. (2024). The impact of past childhood adversity and recent life events on neural responses during fear conditioning. *Journal of Neuroimaging*, *34*(2), 217–223. https://doi.org/10.1111/jon.13174

Katkin, E. S., Wiens, S., & Öhman, A. (2001). Nonconscious Fear Conditioning, Visceral Perception, and the Development of Gut Feelings. *Psychological Science*, *12*(5), 366–370. https://doi.org/10.1111/1467-9280.00368

Kattner, F. (2012). Revisiting the relation between contingency awareness and attention: Evaluative conditioning relies on a contingency focus. *Cognition & Emotion*, *26*(1), 166–175. https://doi.org/10.1080/02699931.2011.565036

Kauvar, I., Richman, E. B., Liu, T. X., Li, C., Vesuna, S., Chibukhchyan, A., Yamada, L., Fogarty, A., Solomon, E., Choi, E. Y., Mortazavi, L., Chau Loo Kung, G., Mukunda, P., Raja, C., Gil-Hernández, D., Patron, K., Zhang, X., Brawer, J., Wrobel, S., … Deisseroth, K. (2025). Conserved brain-wide emergence of emotional response from sensory experience in humans and mice. *Science, 388*(6750), eadt3971. https://doi.org/10.1126/science.adt3971

Kim, S. G., Richter, W., & Uğurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, *37*(4), 631–636. https://doi.org/10.1002/mrm.1910370427

Kirby, L. A. J., & Robinson, J. L. (2017). Affective mapping: An activation likelihood estimation (ALE) meta-analysis. *Brain and Cognition*, *118*, 137–148. https://doi.org/10.1016/j.bandc.2015.04.006

Klucken, T., Kagerer, S., Schweckendiek, J., Tabbert, K., Vaitl, D., & Stark, R. (2009). Neural, electrodermal and behavioral response patterns in contingency aware and unaware subjects during a picture–picture conditioning paradigm. *Neuroscience*, *158*(2), 721–731.

https://doi.org/10.1016/j.neuroscience.2008.09.049

Klucken, T., Schweckendiek, J., Koppe, G., Merz, C. J., Kagerer, S., Walter, B., Sammer, G., Vaitl, D., & Stark, R. (2012). Neural correlates of disgust- and fear-conditioned responses. *Neuroscience*, *201*, 209–218. https://doi.org/10.1016/j.neuroscience.2011.11.007

Knight, D. C., Nguyen, H. T., & Bandettini, P. A. (2003). Expression of conditional fear with and without awareness. *Proceedings of the National Academy of Sciences of the United States of America, 100*(25), 15280–15283. https://doi.org/10.1073/pnas.2535780100

Knight, D. C., Waters, N. S., & Bandettini, P. A. (2009). Neural substrates of explicit and implicit fear memory. *NeuroImage*, *45*(1), 208–214. https://doi.org/10.1016/j.neuroimage.2008.11.015

Kuhn, M., Gerlicher, A. M. V., & Lonsdorf, T. B. (2022). Navigating the manyverse of skin conductance response quantification approaches—A direct comparison of trough-to-peak, baseline correction, and model-based approaches in Ledalab and PsPM. *Psychophysiology*, *59*(9), e14058. https://doi.org/10.1111/psyp.14058

Kyle, B. N., & McNeil, D. W. (2014). Autonomic arousal and experimentally induced pain: A critical review of the literature. *Pain Research & Management*, *19*(3), 159–167. https://doi.org/10.1155/2014/536859

Labrenz, F., Icenhour, A., Benson, S., & Elsenbruch, S. (2015). Contingency Awareness Shapes Acquisition and Extinction of Emotional Responses in a Conditioning Model of Pain-Related Fear. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10.3389/fnbeh.2015.00318

Lam, G. N., Cooper, J., Lipp, O. V., Mayo, L. M., & Ney, L. (2024). Exploration of stress reactivity and fear conditioning on intrusive memory frequency in a conditioned-intrusion paradigm. *Journal of Behavior Therapy and Experimental Psychiatry*, *85*, 101984. https://doi.org/10.1016/j.jbtep.2024.101984

Lambert, E., Treasure, J., Purves, K. L., McGregor, T., Bergou, N., Kan, C., Breen, G., Eley, T. C., & Cardi, V. (2021). Fear conditioning in women with anorexia nervosa and healthy controls: A preliminary study. *Journal of Abnormal Psychology*, *130*(5), 490–497. https://doi.org/10.1037/abn0000549

Landkroon, E., Mertens, G., Sevenster, D., Dibbets, P., & Engelhard, I. M. (2019). Renewal of conditioned fear responses using a film clip as the aversive unconditioned stimulus. *Journal of Behavior Therapy and Experimental Psychiatry*, *65*, 101493. https://doi.org/10.1016/j.jbtep.2019.101493

Leer, A., Smeets, M. A. M., Bulsing, P. J., & van den Hout, M. A. (2011). Odors eliciting fear: A conditioning approach to Idiopathic Environmental Intolerances. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*(2), 240–249. https://doi.org/10.1016/j.jbtep.2010.12.007

Levine, S. M., Kumpf, M., Rupprecht, R., & Schwarzbach, J. V. (2021). Supracategorical fear information revealed by aversively conditioning multiple categories. *Cognitive Neuroscience*, *12*(1), 28–39. https://doi.org/10.1080/17588928.2020.1839039

Li, Y., Menger, N. S., Keil, A., Kosachenko, A. I., Mertens, G., Kotchoubey, B., & Pavlov, Y. G. (2025). Contingency awareness and fear conditioning: A comprehensive examination of associated factors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. https://doi.org/10.1037/xlm0001501

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121–143. https://doi.org/10.1017/S0140525X11000446

Lipp, O. V., Kempnich, C., Jee, S. H., & Arnold, D. H. (2014). Fear Conditioning to Subliminal Fear Relevant and Non Fear Relevant Stimuli. *PLOS ONE*, *9*(9), e99332. https://doi.org/10.1371/journal.pone.0099332

Lissek, S., Baas, J. M. P., Pine, D. S., Orme, K., Dvir, S., Nugent, M., Rosenberger, E., Rawson, E., & Grillon, C. (2005). Airpuff startle probes: An efficacious and less aversive alternative to white-noise. *Biological Psychology*, *68*(3), 283–297. https://doi.org/10.1016/j.biopsycho.2004.07.007

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy*, *43*(11), 1391–1424. https://doi.org/10.1016/j.brat.2004.10.007

Liu, J., Kou, J., Tan, L., Li, H., & Lei, Y. (2025). The complex role of oxytocin in fear acquisition and generalization. *Psychoneuroendocrinology*, *176*, 107421. https://doi.org/10.1016/j.psyneuen.2025.107421

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience & Biobehavioral Reviews*, *80*, 703–728. https://doi.org/10.1016/j.neubiorev.2017.07.007

Lovett-Barron, M., Kaifosh, P., Kheirbek, M. A., Danielson, N., Zaremba, J. D., Reardon, T. R., Turi, G. F., Hen, R., Zemelman, B. V., & Losonczy, A. (2014). Dendritic inhibition in the hippocampus supports fear learning. *Science*, *343*(6173), 857–863. https://doi.org/10.1126/science.1247485

Lovibond, P. F., Liu, J. C. J., Weidemann, G., & Mitchell, C. J. (2011). Awareness is necessary for differential trace and delay eyeblink conditioning in humans. *Biological Psychology*, *87*(3), 393–400. https://doi.org/10.1016/j.biopsycho.2011.05.002

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*(1), 3–26. https://doi.org/10.1037/0097-7403.28.1.3

Lovibond, P. F., & Westbrook, R. F. (2024). The role of expectancy in Pavlovian conditioning. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000516

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, *8*(5), 656–672. https://doi.org/10.1111/j.1469-8986.1971.tb00501.x

Mancinelli, F., Sporrer, J. K., Myrov, V., Melinscak, F., Zimmermann, J., Liu, H., & Bach, D. R. (2024). Dimensionality and optimal combination of autonomic fear-conditioning measures in humans. *Behavior Research Methods*, *56*(6), 6119–6129. https://doi.org/10.3758/s13428-024-02341-3

Manns, J. R., Clark, R. E., & Squire, L. R. (2001). Single-cue delay eyeblink conditioning is unrelated to awareness. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(2), 192–198. https://doi.org/10.3758/CABN.1.2.192

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610. https://doi.org/10.3758/s13428-013-0379-1

Marinkovic, K., Schell, A. M., & Dawson, M. E. (1989). Awareness of the CS-UCS contingency and classical conditioning of skin conductance responses with olfactory CSs. *Biological Psychology*, *29*(1), 39–60. https://doi.org/10.1016/0301-0511(89)90049-5

McDowd, J. M. (2007). An overview of attention: Behavior and brain. *Journal of Neurologic Physical Therapy, 31*(3), 98–103. https://doi.org/10.1097/NPT.0b013e31814d7874

McEchron, M. D., McCabe, P. M., Green, E. J., Llabre, M. M., & Schneiderman, N. (1992). Air puff versus shock unconditioned stimuli in rabbit heart rate conditioning. *Physiology & Behavior*, *51*(1), 195–199. https://doi.org/10.1016/0031-9384(92)90223-o

McGregor, T., Purves, K. L., Barry, T., Constantinou, E., Craske, M. G., Breen, G., Young, K. S., & Eley, T. C. (2023). Introducing the Fear Learning and Anxiety Response (FLARe) app and web portal for the remote delivery of fear conditioning experiments. *Behavior Research Methods, 55*(6), 3164–3178. https://doi.org/10.3758/s13428-022-01952-y

McGregor, T., Purves, K. L., Constantinou, E., Baas, J. M. P., Barry, T. J., Carr, E., Craske, M. G., Lester, K. J., Palaiologou, E., Breen, G., Young, K. S., & Eley, T. C. (2021). Large-scale remote fear conditioning: Demonstration of associations with anxiety using the

FLARe smartphone app. *Depression and Anxiety*, *38*(7), 719–730. https://doi.org/10.1002/da.23146

Mertens, G., Boddez, Y., Krypotos, A.-M., & Engelhard, I. M. (2021). Human fear conditioning is moderated by stimulus contingency instructions. *Biological Psychology*, *158*, 107994. https://doi.org/10.1016/j.biopsycho.2020.107994

Mertens, G., & De Houwer, J. (2016). Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history. *Biological Psychology*, *113*, 91–99. https://doi.org/10.1016/j.biopsycho.2015.11.014

Mertens, G., & Engelhard, I. M. (2020). A systematic review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience & Biobehavioral Reviews*, *108*, 254–268. https://doi.org/10.1016/j.neubiorev.2019.11.012

Meulders, A., Fannes, S., Van Diest, I., De Peuter, S., Vansteenwegen, D., & Van den Bergh, O. (2010). Resistance to extinction in an odor-20% CO_2 inhalation paradigm: Further evidence for a symptom learning account of multiple chemical sensitivity. *Journal of Psychosomatic Research*, *68*(1), 47–56. https://doi.org/10.1016/j.jpsychores.2009.03.009

Meulders, A., Jans, A., & Vlaeyen, J. W. S. (2015). Differences in pain-related fear acquisition and generalization: An experimental study comparing patients with fibromyalgia and healthy controls. *Pain*, *156*(1), 108–122. https://doi.org/10.1016/j.pain.0000000000000016

Meulders, A., Vervliet, B., Fonteyne, R., Baeyens, F., Hermans, D., & Vansteenwegen, D. (2012). Preexposure to (un)predictable shock modulates discriminative fear learning between cue and context: An investigation of the interaction between fear and anxiety. *International Journal of Psychophysiology*, *84*(2), 180–187. https://doi.org/10.1016/j.ijpsycho.2012.02.004

Michalowski, J. M., Pané-Farré, C. A., Löw, A., & Hamm, A. O. (2015). Brain dynamics of visual attention during anticipation and encoding of threat- and safe-cues in spider-phobic individuals. *Social Cognitive and Affective Neuroscience*, *10*(9), 1177–1186. https://doi.org/10.1093/scan/nsv002

Miskovic, V., & Keil, A. (2012). Acquired fears reflected in cortical sensory processing: A review of electrophysiological studies of human classical conditioning. *Psychophysiology, 49*(9), 1230–1241. https://doi.org/10.1111/j.1469-8986.2012.01398.x

Moran, T., Nudler, Y., & Bar-Anan, Y. (2023). Evaluative Conditioning: Past, Present, and Future. *Annual Review of Psychology*, *74*, 245–269. https://doi.org/10.1146/annurev-psych-032420-031815

Morriss, J., Saldarini, F., & van Reekum, C. M. (2019). The role of threat level and intolerance of uncertainty in extinction. *International Journal of Psychophysiology*, *142*, 1–9. https://doi.org/10.1016/j.ijpsycho.2019.05.013

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*(3), 207–233. https://doi.org/10.3758/cabn.3.3.207

Neueder, D., Andreatta, M., & Pauli, P. (2019). Contextual Fear Conditioning and Fear Generalization in Individuals With Panic Attacks. *Frontiers in Behavioral Neuroscience*, *13*. https://doi.org/10.3389/fnbeh.2019.00152

Neumann, D. L., & Waters, A. M. (2006). The use of an unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biological Psychology*, *73*(2), 175–185. https://doi.org/10.1016/j.biopsycho.2006.03.004

Ney, L. J., FitzSimons-Reilly, A., & Lipp, O. V. (2023). Reaction time as an outcome measure during online fear conditioning: Effects of number of trials, age, and levels of processing. *Behaviour Research and Therapy*, *169*, 104406. https://doi.org/10.1016/j.brat.2023.104406

Ney, L. J., Nichols, D. S., & Lipp, O. V. (2023). Fear conditioning depends on the nature of the unconditional stimulus and may be related to hair levels of endocannabinoids. *Psychophysiology*, *60*(8), e14297. https://doi.org/10.1111/psyp.14297

Ney, L. J., O'Donohue, M. P., Lowe, B. G., & Lipp, O. V. (2022). Angry and fearful compared to happy or neutral faces as conditional stimuli in human fear conditioning: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, *139*, 104756. https://doi.org/10.1016/j.neubiorev.2022.104756

Ney, L. J., Schenker, M., & Lipp, O. V. (2022). Combining the trauma film and fear conditioning paradigms: A theoretical review and meta-analysis with relevance to PTSD. *Behaviour Research and Therapy*, *152*, 104081. https://doi.org/10.1016/j.brat.2022.104081

Olejniczak, P. (2006). Neurophysiologic basis of EEG. *Journal of Clinical Neurophysiology*, *23*(3), 186–189. https://doi.org/10.1097/01.wnp.0000220079.61973.6c

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869. https://doi.org/10.1155/2011/156869

Paiva, T. O., Almeida, P. R., Coelho, R. C., Pasion, R., Barbosa, F., Ferreira-Santos, F., Bastos-Leite, A. J., & Marques-Teixeira, J. (2020). The neurophysiological correlates of the triarchic model of psychopathy: An approach to the basic mechanisms of threat conditioning and inhibitory control. *Psychophysiology*, *57*(8), e13567. https://doi.org/10.1111/psyp.13567

Palmer, C. E., Auksztulewicz, R., Ondobaka, S., & Kilner, J. M. (2019). Sensorimotor beta power reflects the precision-weighting afforded to sensory prediction errors. *NeuroImage*, *200*, 59–71. https://doi.org/10.1016/j.neuroimage.2019.06.034

Panitz, C., Hermann, C., & Mueller, E. M. (2015). Conditioned and extinguished fear modulate functional corticocardiac coupling in humans. *Psychophysiology*, *52*(10), 1351–1360. https://doi.org/10.1111/psyp.12498

Panitz, C., Keil, A., & Mueller, E. M. (2019). Extinction-resistant attention to long-term conditioned threat is indexed by selective visuocortical alpha suppression in humans. *Scientific Reports*, *9*(1), 15809. https://doi.org/10.1038/s41598-019-52315-1

Pappens, M., Smets, E., Vansteenwegen, D., Van Den Bergh, O., & Van Diest, I. (2012). Learning to fear suffocation: A new paradigm for interoceptive fear conditioning. *Psychophysiology*, *49*(6), 821–828. https://doi.org/10.1111/j.1469-8986.2012.01357.x

Pappens, M., Van den Bergh, O., Vansteenwegen, D., Ceunen, E., De Peuter, S., & Van Diest, I. (2013). Learning to fear obstructed breathing: Comparing interoceptive and exteroceptive cues. *Biological Psychology*, *92*(1), 36–42. https://doi.org/10.1016/j.biopsycho.2011.05.004

Pastor, M. C., Bradley, M. M., Löw, A., Versace, F., Moltó, J., & Lang, P. J. (2008). Affective picture perception: Emotion, context, and the late positive potential. *Brain Research, 1189*, 145–151. https://doi.org/10.1016/j.brainres.2007.10.072

Pastor, M. C., Rehbein, M. A., Junghöfer, M., Poy, R., López, R., & Moltó, J. (2015). Facing challenges in differential classical conditioning research: Benefits of a hybrid design for simultaneous electrodermal and electroencephalographic recording. *Frontiers in Human Neuroscience, 9*, 336. https://doi.org/10.3389/fnhum.2015.00336

Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford University Press.

Pavlov, Y. G., & Kotchoubey, B. (2019). Classical conditioning in oddball paradigm: A comparison between aversive and name conditioning. *Psychophysiology*, *56*(7), e13370. https://doi.org/10.1111/psyp.13370

Pejic, T., Hermann, A., Vaitl, D., & Stark, R. (2013). Social anxiety modulates amygdala activation during social conditioning. *Social Cognitive and Affective Neuroscience*, *8*(3), 267–276. https://doi.org/10.1093/scan/nsr095

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. https://doi.org/10.1016/s1388-2457(99)00141-8

Pine, D. S., Fyer, A., Grun, J., Phelps, E. A., Szeszko, P. R., Koda, V., Li, W., Ardekani, B., Maguire, E. A., Burgess, N., & Bilder, R. M. (2001). Methods for developmental studies of fear conditioning circuitry. *Biological Psychiatry*, *50*(3), 225–228. https://doi.org/10.1016/s0006-3223(01)01159-3

Plog, E., Antov, M. I., Bierwirth, P., & Stockhorst, U. (2023). Effects of phase synchronization and frequency specificity in the encoding of conditioned fear–a web-based fear conditioning study. *PLOS ONE*, *18*(3), e0281644. https://doi.org/10.1371/journal.pone.0281644

Poplin, T., Ironside, M., Kuplicki, R., Aupperle, R. L., Guinjoan, S. M., Khalsa, S. S., Stewart, J. L., Victor, T. A., Paulus, M. P., & Kirlic, N. (2024). The unique face of comorbid anxiety and depression: Increased frontal, insula and cingulate cortex response during Pavlovian fear-conditioning. *Journal of Affective Disorders*, *366*, 98–105. https://doi.org/10.1016/j.jad.2024.08.143

Pujol, J., Blanco-Hinojo, L., Coronas, R., Esteba-Castillo, S., Rigla, M., Martínez-Vilavella, G., Deus, J., Novell, R., & Caixàs, A. (2018). Mapping the sequence of brain events in response to disgusting food. *Human Brain Mapping*, *39*(1), 369–380. https://doi.org/10.1002/hbm.23848

Purves, K. L., Krebs, G., McGregor, T., Constantinou, E., Lester, K. J., Barry, T. J., Craske, M. G., Young, K. S., Breen, G., & Eley, T. C. (2023). Evidence for distinct genetic and environmental influences on fear acquisition and extinction. *Psychological Medicine*, *53*(3), 1106–1114. https://doi.org/10.1017/S0033291721002580

Raes, A. K., & De Raedt, R. (2011). Interoceptive awareness and unaware fear conditioning: Are subliminal conditioning effects influenced by the manipulation of visceral self-perception? *Consciousness and Cognition*, *20*(4), 1393–1402. https://doi.org/10.1016/j.concog.2011.05.009

Raes, A. K., De Raedt, R., Fias, W., Koster, E. H. W., & Van Damme, S. (2009). Does contingency awareness mediate the influence of emotional learning on the cueing of visual attention? *Psychological Research*, *73*(1), 107–113. https://doi.org/10.1007/s00426-008-0141-y

Raes, A. K., Koster, E. H. W., Van Damme, S., Fias, W., & De Raedt, R. (2010). Aversive Conditioning under Conditions of Restricted Awareness: Effects on Spatial Cueing. *Quarterly Journal of Experimental Psychology*, *63*(12), 2336–2358. https://doi.org/10.1080/17470218.2010.492995

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Raio, C. M., Carmel, D., Carrasco, M., & Phelps, E. A. (2012). Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology*, *22*(12), R477–R479. https://doi.org/10.1016/j.cub.2012.04.023

Rashidi, M., Simon, J. J., Bertsch, K., Wegen, G. V., Ditzen, B., Flor, H., Grinevich, V., Wolf, R. C., & Herpertz, S. C. (2025). Effects of intranasal oxytocin on fear extinction learning. *Neuropsychopharmacology, 50*(3), 548–555. https://doi.org/10.1038/s41386-024-01996-y

Rehbein, M. A., Wessing, I., Zwitserlood, P., Steinberg, C., Eden, A. S., Dobel, C., & Junghöfer, M. (2015). Rapid prefrontal cortex activation towards aversively paired faces and enhanced contingency detection are observed in highly trait-anxious women under challenging conditions. *Frontiers in Behavioral Neuroscience, 9*, 155. https://doi.org/10.3389/fnbeh.2015.00155

Rehman, I., Mahabadi, N., Sanvictores, T., & Rehman, C. I. (2025). Classical Conditioning. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK470326/

Reinhard, J., Mittermeier, A., Brandstetter, L., Mowat, K., Slyschak, A., Reiter, A. M. F., Gamer, M., & Romanos, M. (2024). Fear conditioning and fear generalization in children and adolescents with anxiety disorders. *European Child & Adolescent Psychiatry*, *33*(7), 2163–2172. https://doi.org/10.1007/s00787-023-02304-7

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160206. https://doi.org/10.1098/rstb.2016.0206

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: Social threat induces bodily freeze in humans. *Psychological Science*, *21*(11), 1575–1581. https://doi.org/10.1177/0956797610384746

Sabatinelli, D., Keil, A., Frank, D. W., & Lang, P. J. (2013). Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology*, *92*(3), 513–519. https://doi.org/10.1016/j.biopsycho.2012.04.005

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional perception: Correlation of functional MRI and event-related potentials. *Cerebral Cortex, 17*(5), 1085–1091. https://doi.org/10.1093/cercor/bhl017

San Martín, C., Laborda, M. A., Miguez, G., Sánchez, A., Vervliet, B., & Quezada-Scholz, V. (2023). Relation among, trait anxiety, intolerance to uncertainty and early maltreatment experiences on fear discrimination learning and avoidance generalization online task. *Journal of Behavior Therapy and Experimental Psychiatry*, *81*, 101886. https://doi.org/10.1016/j.jbtep.2023.101886

Sauter, M., Stefani, M., & Mack, W. (2022). Equal Quality for Online and Lab Data: A Direct Comparison from Two Dual-Task Paradigms. *Open Psychology*, *4*(1), 47–59. https://doi.org/10.1515/psych-2022-0003

Schindler, S., Heinemann, J., Bruchmann, M., Moeck, R., & Straube, T. (2022). No trait anxiety influences on early and late differential neuronal responses to aversively conditioned faces across three different tasks. *Cognitive, Affective & Behavioral Neuroscience*, *22*(5), 1157–1171. https://doi.org/10.3758/s13415-022-00998-x

Schlitt, F., Schmidt, K., Merz, C. J., Wolf, O. T., Kleine-Borgmann, J., Elsenbruch, S., Wiech, K., Forkmann, K., & Bingel, U. (2022). Impaired pain-related threat and safety learning in patients with chronic back pain. *Pain*, *163*(8), 1560–1570. https://doi.org/10.1097/j.pain.0000000000002544

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Schultz, D. H., & Helmstetter, F. J. (2010). Classical conditioning of autonomic fear responses is independent of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(4), 495–500. https://doi.org/10.1037/a0020263

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2), 257–261.

Schweckendiek, J., Klucken, T., Merz, C. J., Tabbert, K., Walter, B., Ambach, W., Vaitl, D., & Stark, R. (2011). Weaving the (neuronal) web: Fear learning in spider phobia. *NeuroImage*, *54*(1), 681–688. https://doi.org/10.1016/j.neuroimage.2010.07.049

Seligowski, A. V., Bondy, E., Singleton, P., Orcutt, H. K., Ressler, K. J., & Auerbach, R. P. (2018). Testing neurophysiological markers related to fear-potentiated startle. *Psychiatry Research*, *267*, 195–200. https://doi.org/10.1016/j.psychres.2018.06.023

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

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, *24*(3), 752–775. https://doi.org/10.3758/s13423-016-1170-y

Sheerin, C. M., Moore, A. A., Sawyers, C., Kirkpatrick, R., Hettema, J. M., & Roberson-Nay, R. (2025). Genetic and Environmental Influences on Fear Learning and Generalization. *Psychophysiology*, *62*(4), e70052. https://doi.org/10.1111/psyp.70052

Shner-Livne, G., Barak, N., Shitrit, I., Abend, R., & Shechner, T. (2024). Late positive potential reveals sustained threat contingencies despite extinction in adolescents but not adults. *Psychological Medicine*, *54*(11), 3156–3167. https://doi.org/10.1017/S0033291724001314

Singh, K., Dawson, M. E., Schell, A. M., Courtney, C. G., & Payne, A. F. H. (2013). Can human autonomic classical conditioning occur without contingency awareness? The critical importance of the trial sequence. *Biological Psychology*, *93*(1), 197–205. https://doi.org/10.1016/j.biopsycho.2013.02.007

Sperl, M. F. J., Panitz, C., Hermann, C., & Mueller, E. M. (2016). A pragmatic comparison of noise burst and electric shock unconditioned stimuli for fear conditioning research with many trials: Fear conditioning with noise burst vs. shock as US. *Psychophysiology*, *53*(9), 1352–1365. https://doi.org/10.1111/psyp.12677

Sperl, M. F. J., Wroblewski, A., Mueller, M., Straube, B., & Mueller, E. M. (2021). Learning dynamics of electrophysiological brain signals during human fear conditioning. *NeuroImage*, *226*, 117569. https://doi.org/10.1016/j.neuroimage.2020.117569

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P., & Jacobs, G. S. (1983). Manual for the State-Trait Anxiety Inventory. *Consulting Psychologists Press*.

Staib, M., Abivardi, A., & Bach, D. R. (2020). Primary auditory cortex representation of fear-conditioned musical sounds. *Human Brain Mapping*, *41*(4), 882–891. https://doi.org/10.1002/hbm.24846

Staib, M., & Bach, D. R. (2018). Stimulus-invariant auditory cortex threat encoding during fear conditioning with simple and complex sounds. *NeuroImage*, *166*, 276–284. https://doi.org/10.1016/j.neuroimage.2017.11.009

Stegmann, Y., Andreatta, M., Pauli, P., & Wieser, M. J. (2021). Associative learning shapes visual discrimination in a web-based classical conditioning task. *Scientific Reports*, *11*(1), 15762. https://doi.org/10.1038/s41598-021-95200-6

Stegmann, Y., Teigeler, J., Mirifar, A., Keil, A., & Gamer, M. (2024). Electrocortical Responses in Anticipation of Avoidable and Inevitable Threats: A Multisite Study. *The Journal of Neuroscience*, *44*(42), e0575242024. https://doi.org/10.1523/JNEUROSCI.0575-24.2024

Steinberg, C., Dobel, C., Schupp, H. T., Kissler, J., Elling, L., Pantev, C., & Junghöfer, M. (2012). Rapid and highly resolving: Affective evaluation of olfactorily conditioned faces. *Journal of Cognitive Neuroscience*, *24*(1), 17–27. https://doi.org/10.1162/jocn_a_00067

Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: Evidence for sensory plasticity in early affective perception. *Cerebral Cortex, 16*(6), 876–887. https://doi.org/10.1093/cercor/bhj031

Strand, P. S. (2005). Aversive conditioning. In M. Hersen (Ed.), *Encyclopedia of behavior modification and cognitive behavior therapy* (pp. 684–685). SAGE Publications, Inc. https://doi.org/10.4135/9781412950534

Tabbert, K., Merz, C. J., Klucken, T., Schweckendiek, J., Vaitl, D., Wolf, O. T., & Stark, R. (2011). Influence of contingency awareness on neural, electrodermal and evaluative responses during fear conditioning. *Social Cognitive and Affective Neuroscience*, *6*(4), 495–506. https://doi.org/10.1093/scan/nsq070

Trenado, C., Pedroarena-Leal, N., Cif, L., Nitsche, M., & Ruge, D. (2018). Neural Oscillatory Correlates for Conditioning and Extinction of Fear. *Biomedicines*, *6*(2), 49. https://doi.org/10.3390/biomedicines6020049

Trott, J. M., Hoffman, A. N., Zhuravka, I., & Fanselow, M. S. (2022). Conditional and unconditional components of aversively motivated freezing, flight and darting in mice. *eLife*, *11*, e75663. https://doi.org/10.7554/eLife.75663

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working Memory and Fluid Intelligence: Capacity, Attention Control, and Secondary Memory Retrieval. *Cognitive Psychology*, *71*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Vaitl, D., & Lipp, O. V. (1997). Latent inhibition and autonomic responses: A psychophysiological approach. *Behavioural Brain Research*, *88*(1), 85–93. https://doi.org/10.1016/s0166-4328(97)02310-3

van Ast, V. A., Klumpers, F., Grasman, R. P. P. P., Krypotos, A.-M., & Roelofs, K. (2022). Postural freezing relates to startle potentiation in a human fear-conditioning paradigm. *Psychophysiology*, *59*(4), e13983. https://doi.org/10.1111/psyp.13983

Vollmer, M. (2019). HRVTool: An open-source Matlab toolbox for analyzing heart rate variability. In *2019 Computing in Cardiology (CinC)* (pp. 1–4). IEEE. https://doi.org/10.22489/CinC.2019.032

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, *22*(12), 2864–2885. https://doi.org/10.1162/jocn.2009.21366

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*(1), 1–14. https://doi.org/10.1037/h0069608

Weidemann, G., Satkunarajah, M., & Lovibond, P. F. (2016). I Think, Therefore Eyeblink: The Importance of Contingency Awareness in Conditioning. *Psychological Science*, *27*(4), 467–475. https://doi.org/10.1177/0956797615625973

Wiech, K., & Tracey, I. (2013). Pain, decisions, and actions: A motivational perspective. *Frontiers in Neuroscience*, *7*, 46. https://doi.org/10.3389/fnins.2013.00046

Wiemer, J., Leimeister, F., & Pauli, P. (2021). Subsequent memory effects on event-related potentials in associative fear learning. *Social Cognitive and Affective Neuroscience*, *16*(5), 525–536. https://doi.org/10.1093/scan/nsab015

Wong, P. S., Bernat, E., Snodgrass, M., & Shevrin, H. (2004). Event-related brain correlates of associative learning without awareness. *International Journal of Psychophysiology*, *53*(3), 217–231. https://doi.org/10.1016/j.ijpsycho.2004.04.003

Yin, S., Bo, K., Liu, Y., Thigpen, N., Keil, A., & Ding, M. (2020). Fear conditioning prompts sparser representations of conditioned threat in primary visual cortex. *Social Cognitive and Affective Neuroscience*, *15*(9), 950–964. https://doi.org/10.1093/scan/nsaa122

Yu, K., Beckers, T., Tuerlinckx, F., Vanpaemel, W., & Zaman, J. (2024). The assessment of gender differences in perceptual fear generalization and related processes. *Behaviour*

Research and Therapy, *183*, 104640. https://doi.org/10.1016/j.brat.2024.104640

Zhang, X., Sun, Y., Zhao, C., Yu, F., Li, J., Zhao, Z., & Ai, S. (2025). Transcutaneous vagus nerve stimulation modulates fear memory extinction and neural responses in humans. *Cerebral Cortex, 35*(8), bhaf224. https://doi.org/10.1093/cercor/bhaf224

Zhao, C., Vogel, E., & Awh, E. (2023). Change localization: A highly reliable and sensitive measure of capacity in visual working memory. *Attention, Perception, & Psychophysics*, *85*(5), 1681–1694.

6 Appendices

6.1 Group Homogeneity Checks (Study 1)

Table A1. Checks of homogeneity of the groups.

Group	n1 (Group)	Paired group	n2 (Paired group)	Imbalanced characteristics
Reference	73	Fear ratings	72	smoking within two hours ($\chi^2 = 4.19$, $p = 0.041$)
Reference	73	Expectancy ratings	75	none
Reference	73	Partial reinforcement	75	age ($t = -2.72$, $p = 0.007$) and BAS Fun Seeking ($t = -1.99$, $p = 0.048$)
Reference	73	Trace	75	sex ($\chi^2 = 6.11$, $p = 0.013$), age ($t = -2.87$, $p = 0.005$), BAS Fun Seeking ($t = -2.80$, $p = 0.006$), and BAS Reward Responsiveness($t = -2.43$, $p = 0.016$)
Reference	73	CS preexposure	75	age ($t = -3.44$, $p < 0.001$) and native language ($t = 31.23$, $p < 0.001$)
Reference	73	Story	75	age ($t = -3.68$, $p < 0.001$), native language ($t = 18.45$, $p < 0.001$) and socioeconomic status ($\chi^2 = 4.55$, $p = 0.033$)
Reference	73	Sound CS	75	age ($t = -2.74$, $p = 0.007$) and native language ($\chi^2 = 10.19$, $p = 0.001$)
Reference	73	Sound US	75	age ($t = -2.82$, $p = 0.005$), mind wandering ($z = 2.08$, $p = 0.038$), valence ratings of US ($z = 5.96$, $p < 0.001$), and arousal ratings of US ($z = -3.23$, $p = 0.001$)
SoundUS	75	Overlap CS-US	75	native language ($\chi^2 = 15.43$, $p < 0.001$), handedness ($t = 2.81$, $p = 0.006$), trait anxiety ($t = 2.29$, $p = 0.024$), state anxiety ($t = 2.96$, $p = 0.004$) and IUS ($t = 2.23$, $p = 0.027$)
Fear ratings[a]	72	Unspecific instruction	75	age ($\Delta M = -4.95$, $p = 0.005$), drinking coffee within two hours ($\chi^2 = 5.72$, $p = 0.017$), native language ($\chi^2 = 35.84$, $p < 0.001$), and US arousal ratings ($z = 2.05$, $p = 0.040$)
Fear ratings[a]	72	Specific instruction	75	age ($\Delta M = -5.51$, $p = 0.002$), native language ($\chi^2 = 23.01$, $p < 0.001$), BAS Drive ($\Delta M = -1.06$, $p = 0.042$), and US arousal ratings ($z = 2.38$, $p =$

				0.017)
Unspecific instruction [a]	75	Specific instruction	75	none

Notes. We carried out 12 pairwise comparisons across 21 variables, amounting to a total of 252 tests. Among these tests, 31 (12.3%) yielded statistically significant results. [a]Significant difference in age ($p < 0.001$), drinking coffee within two hours ($p = 0.032$), native language ($p = 0.001$), BAS Drive ($p = 0.038$), and US arousal ratings ($p = 0.037$) was found among groups 2, 11 and 12. Post hoc analyses revealed significant differences between Group 2 and Group 11 on four variables: age, drinking coffee within two hours, native language, and US arousal ratings. Additionally, Group 2 and Group 12 exhibited significant differences in four variables: age, native language, BAS Drive, and US arousal ratings.

6.2 Experimental Task Links (Study 1)

1. Link to the working memory task:

https://gitlab.pavlovia.org/ugpavlov/change_localization_public

https://run.pavlovia.org/ugpavlov/change_localization_public

2. Links to the conditioning task:

Group 1: Reference

https://gitlab.pavlovia.org/Yibo/constudy_noorat_public

https://run.pavlovia.org/Yibo/constudy_noorat_public

Group 2: Online Fear Ratings

https://gitlab.pavlovia.org/Yibo/constudy_fearrating_public

https://run.pavlovia.org/Yibo/constudy_fearrating_public

Group 3: Online Expectancy Ratings

https://gitlab.pavlovia.org/Yibo/constudy_expectancyrating_public

https://run.pavlovia.org/Yibo/constudy_expectancyrating_public

Group 4: 55% Reinforcement Rate

https://gitlab.pavlovia.org/Yibo/constudy_partialreinforcementrate_public

https://run.pavlovia.org/Yibo/constudy_partialreinforcementrate_public

Group 5: Trace Conditioning (2 sec)

https://gitlab.pavlovia.org/Yibo/constudy_trace_public

https://run.pavlovia.org/Yibo/constudy_trace_public

Group 6: CS preexposure

https://gitlab.pavlovia.org/Yibo/constudy_habituation_public

https://run.pavlovia.org/Yibo/constudy_habituation_public

Group 7: Concurrent Task During Acquisition

https://gitlab.pavlovia.org/Yibo/constudy_story_public

https://run.pavlovia.org/Yibo/constudy_story_public

Group 8: Sound CSs

https://gitlab.pavlovia.org/Yibo/constudy_soucs_public

https://run.pavlovia.org/Yibo/constudy_soucs_public

Group 9: Sound US

https://gitlab.pavlovia.org/Yibo/constudy_soundus_public

https://run.pavlovia.org/Yibo/constudy_soundus_public

Group 10: Delay With 500 ms CS/US Overlap

https://gitlab.pavlovia.org/Yibo/constudy_overlappedcs_us_public

https://run.pavlovia.org/Yibo/constudy_overlappedcs_us_public

Group 11: Unspecific Instruction About Contingency

https://gitlab.pavlovia.org/Yibo/constudy_unspecificinstr_public https://run.pavlovia.org/Yibo/constudy_unspecificinstr_public

Group 12: Specific Instruction About Contingency

https://gitlab.pavlovia.org/Yibo/constudy_specificinstr_public

https://run.pavlovia.org/Yibo/constudy_specificinstr_public

6.3 Ratings of Highly Unpleasant Images from the NAPS Database (Study 1)

Table A2. Ratings from the NAPS database of ten highly unpleasant pictures.

Picture	Fear	Disgust	Valence	Arousal
Faces_364_v	4.39 ± 2.09	5.66 ± 1.64	2.14 ± 1.44	6.36 ± 2.46
People_237_h	4.02 ± 2.25	4.93 ± 1.99	2.36 ± 1.45	5.57 ± 2.82
People_220_h	3.95 ± 2.18	5.70 ± 1.59	2.05 ± 1.26	5.91 ± 2.31
People_211_v	3.61 ± 2.17	5.36 ± 1.75	2.43 ± 1.45	5.52 ± 2.51
People_198_h	3.05 ± 2.04	5.71 ± 1.81	2.07 ± 1.13	5.32 ± 2.44
People_216_h	3.73 ± 2.08	4.93 ± 2.05	2.95 ± 1.45	5.61 ± 2.43
People_241_h	3.48 ± 1.96	4.57 ± 1.96	2.59 ± 1.34	4.68 ± 2.08
Faces_365_v	3.77 ± 2.26	4.69 ± 1.81	1.79 ± 1.03	6.31 ± 1.79
Faces_143_v	3.66 ± 2.23	4.59 ± 2.14	1.80 ± 1.12	5.54 ± 2.54
Faces_149_v	3.67 ± 2.12	3.69 ± 1.89	2.33 ± 1.36	5.85 ± 1.89

6.4 Computational Model: Estimating the Number of Extinction Trials (Study 1)

MATLAB code for Calculating the Number of Extinction Trials

```
%%% START MATLAB CODE %%%%%%%%%%%%%
% learn weights w based on inputs u and rewards r with learning rate eta
eta = 0.4;
u= 1;
r = [ 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0]';
```

```
trials = size(r,1); % establish number of trials
dim = size(u,2); % establish number of feature dimensions
w = zeros(trials+1, dim);
delta = zeros(trials, dim);

for t=2:trials
   delta(t) = u .* eta .* (r(t)-w(t));
   w(t+1) = w(t) + delta(t);
end

w(isinf(w)) = 1;
w = w(2:end);
%%% END MATLAB CODE %%%%%%%%%%%%%
```

6.5 Questions for the Audio Story Task (Group 7, Study 1)

The audio story presented to the participants is based on a short story Lazy Jack (available here: https://www.storynory.com/lazy-jack/). Before exclusion, ninety participants answered an average of 3.50 questions correctly with a standard deviation of 1.46.

1. What did Jack's mother warn him about if he didn't start working?
   Option 1: She would turn him into a donkey.
   Option 2: She would send him to live with his grandmother.
   Option 3: She would kick him out of the house.
   Option 4: She would take away his favorite toys.
2. How did Jack carry the jar of milk on his way back home?
   Option 1: In his jacket pocket.
   Option 2: On his head.
   Option 3: Tied to a string.
   Option 4: In his hands.
3. Why did Jack end up losing the cream cheese given by the farmer?
   Option 1: He ate it on the way home.
   Option 2: He carried it on his head.
   Option 3: It got spoiled in his pocket.
   Option 4: He was attacked by a cat.
4. How did Jack try to carry the tomcat given by the baker?
   Option 1: Tied to a string and dragged along.
   Option 2: In his hands.
   Option 3: On his head.

Option 4: On his shoulders.

5. What happened to the rich man's daughter after she laughed at Jack?
    - Option 1: She was attacked by a cat.
    - Option 2: She fell in love with Jack at first sight.
    - Option 3: She turned into a donkey.
    - Option 4: She regained her hearing and speech.

6.6 Comparison of US Affective Ratings Between Groups (Study 1)

Table A3. US affective ratings comparison between the groups.

Group	USvalence	USarousal	Paired group	*t* (USvalence)	*p* (USvalence)	*t* (USarousal)	*p* (USarousal)
Reference	15.92 ± 14.81	60.22 ± 26.86					
Fear ratings	17.02 ± 15.64	64.69 ± 25.27	Reference	0.44	0.662	1.03	0.304
Expectancy ratings	18.84 ± 16.86	57.43 ± 27.91	Reference	1.12	0.264	-0.62	0.535
Partial reinforcement	19.48 ± 16.96	61.07 ± 25.90	Reference	1.36	0.175	0.19	0.846
Trace	17.52 ± 10.58	61.89 ± 23.41	Reference	0.75	0.452	0.40	0.688
CS preexposure	14.47 ± 11.84	61.80 ± 27.45	Reference	-0.66	0.513	0.35	0.725
Story	18.52 ± 13.46	56.21 ± 28.42	Reference	1.12	0.265	-0.88	0.379
Sound CS	18.43 ± 16.05	62.07 ± 25.34	Reference	0.99	0.323	0.43	0.667
Sound US	5.45 ± 11.08	72.58 ± 30.19	Reference	-4.86	**<0.001**	2.63	**0.009**
Overlap CS-US	6.13 ± 11.40	75.36 ± 29.81	Sound US	0.37	0.709	0.57	0.571
Unspecific instruction	18.67 ± 14.81	55.76 ± 27.84	Fear ratings	0.66	0.513	-2.04	**0.043**

Specific instruction	19.08 ± 16.22	53.46 ± 31.10	Fear ratings	0.78	0.435	-2.41	**0.017**
			Unspecific instruction	0.16	0.872	-0.48	0.634

Notes. US = unconditioned stimuli; CS = conditioned stimuli.

6.7 Comparison of Characteristics Between Aware and Unaware Participants (Excluding Loud US Groups, Study 1)

Table A4. Comparison of Characteristics between unaware (n = 237) and aware (n = 508) participants without the groups with loud noise as the US.

Variables	Unaware Subjects (n = 237)	Aware Subjects (n = 508)	χ^2 / t / Z	*p*	Effect size (*V/d/r*)
Sex			0.13	0.719	0.013
Female	98/237 (41.4%)	203/508 (40.0%)			
Male	139/237 (58.6%)	305/508 (60.0%)			
Native Language			5.57	0.018	0.086
Others	111/237 (46.8%)	285/508 (56.1%)			
English	126/237 (53.2%)	223/508 (43.9%)			
Education			2.07	0.151	0.053
Below bachelor	74/237 (31.2%)	186/508 (36.6%)			
Bachelor or higher	163/237 (68.8%)	322/508 (63.4%)			
Alcohol[a]			-	0.006	0.108
No	229/237 (96.6%)	505/508 (99.4%)			
Yes	8/237 (3.4%)	3/508 (0.6%)			
Smoking			0.99	0.320	0.036
No	209/237 (88.2%)	460/508 (90.6%)			
Yes	28/237 (11.8%)	48/508 (9.4%)			
Coffee			0.28	0.596	0.019
No	153/237 (64.6%)	338/508 (66.5%)			
Yes	84/237 (35.4%)	170/508 (33.5%)			
Other drugs[a]			-	0.318	0.054
No	236/237 (99.6%)	508/508 (100.0%)			
Yes	1/237 (0.4%)	0/508 (0.0%)			
Socioeconomic			0.03	0.866	0.006
Below 6th rung	116/237 (48.9%)	252/508 (49.6%)			
6th rung or higher	121/237 (51.1%)	256/508 (50.4%)			

Age	34.30 ± 10.28	32.32 ± 9.68	2.55	0.011	0.201
Handedness	42.93 ± 10.35	43.95 ± 9.19	-1.36	0.175	-0.107
STAI_trait	42.59 ± 11.28	43.28 ± 13.22	-0.69	0.488	-0.055
STAI_state	37.50 ± 11.66	35.87 ± 11.98	1.75	0.081	0.137
IUS	71.52 ± 21.29	71.57 ± 20.49	-0.03	0.974	-0.003
BAS Drive	8.97 ± 2.54	9.26 ± 2.63	-1.45	0.146	-0.114
BAS Fun Seeking	8.78 ± 2.39	8.88 ± 2.25	-0.59	0.552	-0.047
BAS Reward Responsiveness	8.64 ± 2.94	8.25 ± 2.76	1.76	0.078	0.139
BIS	15.04 ± 4.29	14.33 ± 4.04	2.18	0.029	0.172
Working memory accuracy[b]	0.60 ± 0.13	0.62 ± 0.12	-1.46	0.144	-0.115
Mind wandering[b]	1.00 (0.00, 2.00)	1.00 (0.00, 2.00)	0.55	0.560	0.020
Valence ratings of US	16.52 (5.50, 26.56)	16.45 (5.56, 25.48)	0.16	0.875	0.006
Arousal ratings of US	63.58 (39.48, 80.88)	62.93 (45.22, 81.11)	-0.30	0.767	-0.011

Notes. US = unconditioned stimuli; STAI = State-Trait Anxiety Inventory; IUS = Intolerance of Uncertainty Scale; BAS = behavioral activation system; BIS = behavioral inhibition system; - = no test statistic is available. [a] The variable was analyzed by Fisher's exact test (the same as below). Mind wandering, valence ratings of US, and arousal ratings of US were reported as the median (first quartile, third quartile), and the Mann-Whitney U test was used for the analyses. The means of mind wandering were 1.44 (SD = 1.73) and 1.34 (SD = 1.61) for unaware and aware participants, respectively. The means of US valence ratings were 17.97 (SD = 14.76) and 17.73 (SD = 14.90) for unaware and aware participants, respectively. The means of US arousal ratings were 58.52 (SD = 28.50) and 59.87 (SD = 26.41) for unaware and aware participants, respectively. [b] For the working memory accuracy and mind wandering, two participants had incomplete data, so the number of participants included in the analyses was 743.

6.8 Participant Characteristics by Type of Unconditioned Stimulus (Study 2)

Table A5. Demographic and psychological characteristics across airpuff, image, and electric shock US groups.

Variables	US airpuff (n = 35)	US image (n = 52)	US shock (n = 28)	p
Sex				0.619
Female	24/35 (68.6%)	36/52 (69.2%)	22/28 (78.6%)	
Male	11/35 (31.4%)	16/52 (30.8%)	6/28 (21.4%)	
Native Language				0.361
Others	31/35 (88.6%)	50/52 (96.2%)	26/28 (92.9%)	

English	4/35 (11.4%)	2/52 (3.8%)	2/28 (7.1%)	
Education				0.198
Below bachelor	9/35 (25.7%)	16/52 (30.8%)	13/28 (46.4%)	
Bachelor or higher	26/35 (74.3%)	36/52 (69.2%)	15/28 (53.6%)	
Alcohol				1.000
No	35/35 (100.0%)	51/52 (98.1%)	28/28 (100.0%)	
Yes	0/35 (0.0%)	1/52 (1.9%)	6/28 (0.0%)	
Smoking				0.116
No	35/35 (100.0%)	50/52 (96.2%)	25/28 (89.3%)	
Yes	0/35 (0.0%)	2/52 (3.8%)	3/28 (10.7%)	
Coffee				0.174
No	32/35 (91.4%)	41/52 (78.8%)	21/28 (75.0%)	
Yes	3/35 (8.6%)	11/52 (21.2%)	7/28 (25.0%)	
Other drugs				1.000
No	35/35 (100.0%)	51/52 (98.1%)	28/28 (100.0%)	
Yes	0/35 (0.0%)	1/52 (1.9%)	0/28 (0.0%)	
Socioeconomic				0.352
Below 6th rung	31/35 (88.6%)	42/52 (80.8%)	21/28 (75.0%)	
6th rung or higher	4/35 (11.4%)	10/52 (19.2%)	7/28 (25.0%)	
Age	25.09±4.20	25.98±5.28	24.00±3.43	0.182
Handedness	43.29±8.38	45.13±6.10	42.39±10.00	0.290
STAI_trait	37.06±7.86	44.56±10.36	44.39±10.97	**0.001***
STAI_state	31.46±7.82	34.65±10.79	36.00±11.90	0.187
IUS	59.83±13.74	66.69±17.93	67.68±18.84	0.114
BAS Drive	8.20±2.25	8.21±2.15	8.68±2.29	0.619
BAS Fun Seeking	8.14±2.82	7.52±2.03	7.75±2.52	0.499
BAS Reward Responsiveness	8.43±3.42	7.38±2.32	7.68±2.18	0.202
BIS	15.57±3.37	13.88±3.84	13.54±3.54	**0.049***

Notes. BIS = Behavioral Inhibition System; BAS = Behavioral Activation System; STAI = State-Trait Anxiety Inventory; IUS = Intolerance of Uncertainty Scale. Categorical variables are presented as frequencies and percentages, while continuous variables are expressed as mean ± standard deviation. Group comparisons for categorical variables were conducted using the Chi-square test for Sex and Education, and Fisher's exact test for all other categorical variables. For continuous variables, one-way analysis of variance (ANOVA) was used to assess group differences. * *p* < .05.

6.9 US Valence and Arousal Ratings Comparison between the Shock, Airpuff, and Image US Groups (Image Group Ratings Based on the First Aversive Image Only)

Table A6. Differences in valence and arousal ratings across the shock, airpuff, and image US groups (image group ratings based on the first aversive image only)

	Images ($n = 51$)	Shocks ($n = 28$)	Airpuffs ($n = 34$)	Kruskal-Wallis $\chi^2(2)$	*p*	η^2
Valence	22.48 ± 22.73	32.99 ± 22.85	40.36 ± 23.76	12.07	**0.002***	0.09
Arousal	57.98 ± 32.28	47.54 ± 31.38	42.29 ± 30.75	5.68	0.058	0.03

Notes. Values represent mean ± standard deviation. **p* < .05.