Aus der Klinik für Psychosomatik und Psychotherapeutische Medizin

des Zentralinstituts für Seelische Gesundheit

der Medizinischen Fakultät Mannheim

der Universität Heidelberg

(Direktor: Prof. Dr. med. Christian Schmahl)

The influence of contextual threat and safety transmitted by social learning on visual working memory and item/source memory and its modulation by adverse childhood experiences and (social) anxiety

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum (Dr.sc.hum.)

der

Medizinischen Fakultät Mannheim

der Ruprecht-Karls-Universität zu Heidelberg

vorgelegt von

Sabine Schellhaas

aus

Bad Friedrichshall

2022

Dekan: Prof. Dr. med. Sergij Goerdt

Referent: Prof. Dr. Christian Schmahl

Table of Contents

General Thesis Discussion

Abbreviations

ABC	Arousal Biased Competition
ACC	Anterior Cingulate Cortex
ACE	Adverse Childhood Experiences
ACT	Attentional Control Theory
ANOVA	Analysis of Variance
BDI II	Beck Depression Inventory II
CDA	Contra-lateral Delay Activity
CR	Conditioned Response
CS	Conditioned Stimulus
(A)CSIM	(Average) Conditional Source Identification Measure
CTQ	Childhood Trauma Questionnaire
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders, 4th edition, text revision
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th edition
EDA	Electrodermal Activity
EEG	Electroencephalogram
EMG	Electromyography
ERP	Event-Related Potential
FAR	False Alarm Rate
FNEK	Brief Fear of Negative Evaluation Scale
HR	Hit Rate

ISM	Item/Source Memory
LPP	ERP component; Late Positive Potential
MPT	Multinomial Processing Tree (model)
N170	ERP component
RT	Reaction Time
SCID-I	Structured Clinical Interview for DSM–IV Axis I Disorders
SCID-5	Structured Interview for DSM–5 Personality Disorders
SCR	Skin Conductance Response
SPIN	Social Phobia Inventory
STAI-S/T	State/Trait Anxiety Inventory
ToS	Threat-of-Shock Paradigm
UR	Unconditioned Response
US	Unconditioned Stimulus
VWM	Visual Working Memory

Chapter I

Theoretical Background

1.1 Memory

Well-functioning memory processes contribute to adaptive (social) behaviors in everyday life. The human memory is pretty good, but far from perfect. “I know or I think I know this person” – everyone has had these thoughts when encountering a vaguely familiar person. Depending on the importance of the situation, failing to recognize a person or acquaintance can result in discomfort from an unsuccessful social interaction, or in dangerous situations, result in missing a possible threat (Baddeley et al., 2009). When memory failures occur systematically and frequently, adequate cognitive and social functioning becomes strenuous. Therefore, understanding the underlying mechanisms of adaptive and maladaptive memory functioning is important for preventing memory dysfunctionality in terms of pathological distortions.

Memory systems. Memory is a multifaceted processing system with multiple memory categories (Ruchkin et al., 1992). Following an information-processing approach, the environment provides initial information, which is processed through sensory memory via attentional mechanisms (Crowder, 1978; Öğmen & Herzog, 2016; Atkinson & Shiffrin, 1968). Once the information has gained access to the organism, it is further processed in working memory. Working memory refers to a limited-capacity brain system that provides temporary storage and manipulation of the provided information necessary for higher cognitive tasks such as learning, problem solving, language comprehension (Baddeley, 1992; Baddeley, 2007). Through consolidation processes (Squire et al., 2015), the information gains access to long-term memory which consists of an explicit memory part (i.e., declarative, such as episodic and semantic memory) and an implicit memory part (i.e., non-declarative, such as conditioning and skills). Episodic memory refers to a system that enables to remember specific events in one’s life (Baddeley et al., 2009). Various factors can modulate memory processing and different stages of the information-processing stream, such as dynamic context conditions, timing, attention, and emotion. In three studies, this thesis focuses on a comparison of the short-term processing of information in visual working memory and the process of recognition in the form of long-term storage of item and source (context) information. Additionally, the detrimental or beneficial influence of differential threat and safety contexts on these cognitive systems is assessed. Associations to individuals' expression of anxiety and exposure to early stress experiences are also investigated.

First, visual working memory refers to the system that temporarily maintains and stores visually transmitted information. Working memory includes central cognitive processes that are underlying many higher-level cognitive functions, such as problem solving, executive functioning, learning and memory (Baddeley, 2003; van Ast et al., 2014). The working memory system seems to be domain specific (Engle, 2001; Cocchini et al., 2002; Fougnie et al., 2015). Different modalities (e.g., verbal, visual, auditory) are found to compete within the same domain but less so across different domains in dual-task paradigms (Gruber, 2001; Piccardi et al., 2015). One domain specializes in the maintenance of affective stimuli, contributing to affective working memory, which interferes with cognitive functioning and processing (Mikels et al., 2008; Mikels, & Reuter-Lorenz, 2019). Additionally, working memory capacity in general and domain capacity is limited and varies with regard to interindividual differences (e.g., fluid intelligence, Engle et al., 1999, van Ast et al., 2014). Working memory capacity has been found to be reduced in several mental disorders such as schizophrenia and depression (e.g., Barch & Smith, 2008; Pelosi et al., 2000) and when an individual is under stress or arousal (e.g., psychosocial stress, Schoofs et al., 2008). As its capacity is limited, distraction and interferences are possible and only around 3-4 objects can be maintained simultaneously (Engle, 2001; Vogel et al., 2005).

One way to measure visual working memory capacity is through change detection. In a change detection task, individuals monitor an array of stimuli for a very short amount of time

(200-300 ms) and must identify changes in the presentation of these stimuli (Pessoa & Ungerleider, 2004; Sessa et al. 2011; Stout et al., 2013). Rapid change perception is closely related to attentional processes (Simons & Rensink, 2005). For instance, individuals perform rather poorly at detecting a change in these brief time windows without attentional focus on the change at hand (Beck et al., 2001). Change detection is associated with the activation of a brain network consisting of parietal and frontal brain regions with involvement of the pulvinar, the cerebellum and the inferior temporal gyrus (Pesoa & Ungerleider, 2004). Furthermore, change detection is well suited to investigate the temporal underpinnings of visual working memory capacity in the form of event-related potentials in the electroencephalogram, e.g., contra-lateral delay activity (CDA; Stout et al., 2013; Ikkai et al., 2010; Luck et al., 2000; Luck & Vogel,1997).

The second system, item/source memory, refers to the conditions and contexts (i.e., origins) under which an information is initially acquired (Doerksen & Shimamura, 2001). This includes the temporal and spatial factors that surround a piece of information as well as the social conditions and the modality through which the information was perceived (Wilding, 1999; May et al., 2005). Source memory is a part of episodic memory and identifying the source of a remembered piece of information is important for many cognitive tasks, be they in laboratory settings or in real-life. Remembering the source increases the livelihood and subjective certainty of a memory and it also provides additional cues for retrieval (Johnson et al., 1997). Moreover, source memory shows how detailed memories are stored and ultimately how well they are learned (Ventura-Bort et al., 2016a). The source memory framework by Johnson et al. (1993) suggests, that the source information is not saved as a memory tag together with the central information, but meaningful associations must be regained and reconstructed and then attributed to the most probable source. For this, not only the quality of encoding situations but also the quality of recognition matters and factors such as time pressure, severe stress, and (emotional) distraction, can lead to impaired source reconstruction. The initial acquisition of memory occurs through learning processes (i.e., encoding of environmental stimuli). Available surrounding information is perceived or processed either externally through interaction with the environment or internally via introspection (Klein, 2015).

Behavioral performance measures of memory. Memory performance can be evaluated by using standard measures derived from signal detection theory (Stanislaw & Todorov, 1999). These measures can be used in case when processing categorical data and when making a decision based on the differentiation of signal and noise (distractor). Task hit rates (HR, correct responses divided by all responses in target trials) and false alarm rates (FAR, incorrect responses divided by all responses in non-target trials) were calculated for all memory tasks in all three studies. For the item/source memory task in all studies, item recognition (HR-FAR) and a measure of source identification (average conditional source identification measure [AC-SIM], which is the number of all correct source identifications divided by the overall number of items that were correctly identified as targets; Bell & Buchner, 2011; Bröder & Meiser, 2007) were also included. Visual working memory in Study 2 was quantified with a standard index of sensitivity, reflecting change detection ability (d’ = ZHR-ZFA [inverse-normal transform]) and an index of working memory capacity (i.e., the number of encoded items in each condition [Cowan’s k (k = load*(HR-FA)]).

“Traditional” memory performance measures have several shortcomings when evaluating performance in an item/source memory task when based on the assumptions of the underlying source monitoring framework (Johnson et al., 1993; Bröder & Meiser, 2007). The main reason for using an additional model-based approach to evaluate the behavioral performance in the item/source memory task is that traditional measures focus mostly on memory processes. However, source attribution can involve a variety of contributing factors, such as biases, metacognitive awareness, plausibility beliefs, but also current motivations and intentions (source monitoring framework, Johnson et al., 1993). The two main cognitive processes that play intro source monitoring are memory processes (item and source memory) and guessing (i.e., biases). Guessing, however, is not understood as a random act of source attribution but rather as the best educated guess by an individual based on assumptions and known qualities of the source(s).

Multinomial processing tree (MPT) modeling aims at disentangling memory from guessing processes (Batchelder & Riefer, 1990; Bröder & Meiser, 2007; Bayen et al., 1996). The models are also useful for unraveling latent cognitive processes and modeling explicit event sequences. Including Bayesian hierarchical models within the MPTs (instead of maximum-likelihood parameter estimation; Read & Cressie, 1988) can account for heterogeneity between individuals by estimating associations between model parameters and other data (e.g., anxiety scores; Klauer, 2010; Arnold, Bayen, & Böhm, 2015). MPT models are based on categorical data which are assumed to follow an asymptotic multivariate normal distribution (Erdfelder et al., 2009). The underlying cognitive processes are modeled as parameter estimates (i.e., probabilities) of various latent cognitive states (i.e., unobservable), namely item and source memory and item and source guessing. These probabilities for each source are modeled by processing trees (Erdfelder et al., 2009; see Figure 2 and Chapter II 2.3 for more detailed information). The two-high-threshold source memory (2HTSM) model of Bayen et al. (1996) was the basis for all three studies in this thesis. This model had been tested to the design and paradigms that were used in this thesis with good model fit and reliable parameter estimates and was therefore suitable for the current research questions (Arnold et al., 2021), especially in the context of inter-individual differences, such as depression and anxiety (Arnold, Bayen, & Böhm, 2015).

Neural and psychophysiological measures of memory. Besides measures at the behavioral level, event-related potentials (ERPs) were used. These were derived from Electroencephalography (EEG) as a measure to understand the neural processes underlying complex memory processes. Measures for autonomic and somatic nervous system activation served as indicators of defensive behavior (Electrodermal activity [EDA], and startle electromyography [EMG]).

Mental systems can be understood via cognitive operations but also through neural activity (Posner & Petersen, 1990; Petersen & Posner, 2012; Schupp et al., 2006; Kok, 2001). In this context, the ERPs contained in the EEG signal provide psychophysiological markers for memory, attention, and perceptual processes (Schupp et al., 2006; Chun & Turk-Browne, 2007; Friedman & Johnson, 2000) and are excellent means following the time course of stimulus processing (Luck & Kappenman, 2012). In the context of attention and memory research, several early and late components are studied to be associated with encoding, retrieval (old/new effect; early and late positivity, e.g., Weymar et al., 2013) and threat effects (P1, P2, LPP; Bublatzky & Schupp, 2012).

For memory, several studies report an association with the P300 component, which is thought to indicate the process of updating new representations in working memory and signal the limited processing capacity of working memory (e.g., Donchin, 1981; Hyun et al., 2009; Awh et al., 2006; Kok, 2001). In addition, CDA (Stout et al., 2013) or sustained posterior contralateral negativity (SPCN; Sessa et al., 2011) are suggested as indicators of the maintenance of working memory. A further, face-specific (i.e., the stimuli of interest in this thesis), potential in EEG is an early late negative potential (N170; Schindler & Bublatzky, 2020; Morgan et al., 2008). It is suggested to vary with different mental states and contextual features an individual is confronted with and can therefore be a valuable marker for the perception of threat (Schupp et al., 2004). Also, altered processing due to individual differences but also maladaptive factors leading to or stemming from pathological alterations (i.e., anxiety, trauma) can be made visible with the ERP technique (Hajcak et al., 2019) and may provide markers for biased memory processing for certain cognitive tasks and mental disorders. These ERP measures were used in Studies 1 and 2.

Because both memory systems described above are central to the performance of complex cognitive tasks, such as decision making and behavior planning (Johnson et al., 1997), it is important to explore the underlying mechanisms and influential factors of memory. One of the widely researched factors that influences memory is emotional arousal (Eysenck, 1976; Clark, Milberg, & Ross, 1983; Bradley et al., 1992).

1.2 Impact of Anxiety and Arousal on Memory

Contextual factors, such as the presence of a threat, influence attentional and neural processing (Bublatzky & Schupp, 2012, Wieser & Brosch, 2012). This directly affects the storage and retrieval of relevant memory content (Schwabe et al., 2012). There is a long research tradition investigating the effect of affective arousal and/or anxiety on memory processes with mixed results, ranging from positive effects of arousal on memory performance to no impact and to detrimental effects on memory and attention (Dolcos et al., 2020; Robinson et al., 2013). Various factors play into the association of arousal on memory depending for instance on the memory modality, the nature of the task and whether the arousing events occur during learning, consolidation, or retrieval/recognition. Most approaches therefore include the factor of competing resources, where opposing effects of differential task demands (i.e., cognitive load) can be found on the (non)disruptive effect of emotion. This depends also on the type of information that is being processed (e.g., verbal or spatial; Vytal et al., 2012). There is evidence that arousing events are more likely to be remembered (Kensinger, 2009). During learning (i.e., encoding), arousal-eliciting events are more easily detected and attended to (Dolan & Vuilleumier, 2003) and less easily dismissed during consolidation processes (LaBar & Phelps, 1998). This becomes evident by stronger enhancement of the memory effect through arousal over longer retention intervals.

The release of stress hormones is assumed to play a crucial role in the association of enhanced memory and emotional arousal (Roozendaal et al., 2009). Stress disrupts the normal functioning of an individual and stems from an interaction between environmental and individual factors that is closely linked to anxiety, arousal, and cognitive load (Tepas & Price, 2000;

McGrath, 1976). Stress activates the hypothalamic-pituitary-adrenal (HPA) axis and the hypothalamus releases corticotropin-releasing-hormone (CRH) which in turn leads to the release of adrenocorticotropin (ACTH), another hormone, from the pituitary gland, located in the brain. ACTH triggers the secretion of the stress hormones, the glucocorticoids (e.g., cortisol) and the catecholamines (adrenaline and noradrenaline) from the adrenal glands, above the kidney (Lupien et al., 2007).

The hippocampus, the amygdala and the frontal lobes contain many receptors for glucocorticoids and are highly involved in learning and memory mechanisms. The connectivity between active regions responsible for emotion processing (e.g., amygdala and orbitofrontal cortex) and areas linked to perceptual processing and consolidation (e.g., prefrontal cortex) is strengthened. The storage of arousing information is thereby facilitated through the involvement of this large brain network, even for very short periods of arousal in a rapidly changing environment (Anderson et al., 2006). When encoded in an emotionally negative context, recall is predicted by activation in the amygdala (Erk et al., 2003), in contrast to an encoding in an emotionally positive context where the right anterior parahippocampal and extrastriate visual brain areas are active. Thus, context matters as well as the type of stimulus (e.g., faces vs scenes; Keightley et al., 2010). For instance, for scenes additional visual processing was predictive of recognition whereas faces seem to rely more on cognitive control mediated by rostrolateral prefrontal regions (based on correct recognition performance).

Overall, the emotional memory-enhancement effect is, however, less conclusive when it comes to retrieval and recognition processes and the interaction between emotional arousal, attention and memory is complex (Dolcos et al., 2020). It remains unclear which details are remembered of the arousing event. Some studies suggest that the enhancements happen focally and not peripherally, concentrating on the information that elicits arousal and thus ignoring the surrounding context (Mather, 2007). By instructing a participant to remember all aspects of the scene, the central stimulus and the context can be bound together leading to improved memory performance (Kensinger et al., 2005). Other studies suggest that when an emotional arousing stimulus is task-relevant, it leads to memory enhancement, if an arousing stimulus, however, is task-irrelevant it leads to detrimental effects (prioritization of arousing information; Robinson et al., 2013; Bar-Haim et al., 2007).

Different theoretical models try to integrate and to explain these opposing effects. The Arousal Biased Competition model (ABC model; Mather & Sutherland, 2011) assumes that the processing and memory of arousing, and therefore salient stimuli is enhanced when attention must be selected and focused. The underlying assumption is that multiple items in the visual field compete for neural representation (Bundesen et al., 2005; Deco & Rolls, 2005; Desimone, 1998; Desimone & Duncan, 1995; Kastner & Ungerleider, 2001; Miller & Cohen, 2001). This beneficial effect of arousal starts at basic perception processes and continues into consolidation of long-term storage. Less salient stimuli must compete more vigorously for attentional resources resulting in a weakened processing (Sutherland & Mather, 2012). As competition starts in the visual cortex and is integrated in different brain areas, a superiorly attended item will also be more likely gain access to prefrontal and parietal cortices (Lee et al., 2014).

Salience can be driven by perceptual features of a stimulus that are unexpected or new, or by motivational and goal-driven internal processes. Therefore, the emotional relevance of a top-down acquired, threatening quality of a stimulus should lead to a shift in attention towards the threat. Attentional Control Theory (ACT; Eysenck et al., 2007) expands on the note that arousal promotes stimulus- (threat-)driven attention and impairs goal-driven motivational focus by reducing control over the attention processes. Compensating strategies can be used to improve efficiency in task performance when the quality of performance would have otherwise suffered. In this way, arousal impairs task performance for task-irrelevant stimuli only when there are insufficient cognitive resources to counterbalance its detrimental effects on cognition, thereby reducing inhibition ability (Eysenck, & Derakshan, 2011). Deleterious effects of arousal on attention control are highest when carrying out highly demanding tasks and when the presence of a possible threat requires maintaining high levels of vigilance (Mathews, 1990).

In summary, the effect of arousal (i.e., threat) as a contextual factor on cognitive performance, especially when it comes to processing and remembering of neutral items, is complex and depends on various factors, including memory modality and inter-individual differences.

1.3 Social Threat Learning

The introduction of threat as an environmental factor is possible in many ways. One of the most prominent models of threat learning is Pavlovian fear conditioning (Pavlov, 1927, 2010; Davis 1992; Fendt & Fanselow 1999; LeDoux 2000; Maren, 2001). Fear conditioning is established by associating an aversive unconditioned stimulus (US) with a perceptual representation (e.g., visual) of an conditioned stimulus (CS). The repeated pairing of the CS with the US leads to the formation of a conditioned response (CR) which is similar to the unconditioned response (UR). However, human behavior and the processing of stimuli are not only driven by direct experience. They can also be transmitted vicariously through observation or verbal instruction using social threat learning (Bandura, & Walters, 1977; Laland, 2004; Heyes, 2012; Olsson, & Phelps, 2007). For example, pronounced fear and avoidance of a certain place can be caused both by having been robbed at that place in the past or by having witnessed a robbery there. Finally, a verbal warning that there is a high incidence of robbery at a given location can be sufficient for this location to be avoided or even associated with fear. Thus, people learn both directly through their own experiences (e.g., Pavlovian conditioning) and indirectly through social interactions, that is, through observation or communication (Olsson & Phelps, 2007; Rachman, 1977; Haaker et al., 2017; Robinson et al., 2013).

Different defensive behaviors are activated depending on the proximity or imminence of threat (Dolcos et al., 2020; Hamm, 2020; Fanselow, 1994). The predator imminence model (Fanselow & Lester, 1988; Hoffman et al., 2022) suggests that defensive behavior is organized on a predator imminence continuum. If a threat is perceived as probable in a given environment, informed by prior experience or by external verbal communication, then pre-encounter defense mechanisms are activated leading to an unspecific heightened vigilance and to cessation of appetitive behavior. Autonomic arousal increases and feelings of anxiety, apprehension and worry are reported, especially if the threat is novel and ambiguous (Davis et al., 2010). When the threat-signaling cue becomes evident but is still far away, post-encounter defensive mechanisms are activated, leading to heart rate deceleration, enhanced startle reflex, attentive motor freezing, and increasing feelings of fear. When threat is imminent (circa strike defense), fight/flight responses are initiated. When fight/flight responses are not possible, tonic immobilization occurs.

In healthy individuals, both observational learning and instructional threat learning lead to rapid acquisition of fear responses, comparable to conditioning (Olsson & Phelps, 2007; Öhman & Mineka, 2001; Debiec & Olsson, 2017). In direct fear conditioning, the CS-US association is formed by the connectivity of different brain areas (amygdala, sensory cortices, thalamus and regions associated with pain processing, anterior cingulate cortex [ACC], insular cortex). Observational and verbal fear learning are assumed to share similar neural mechanisms with a few exceptions. For observational learning, a distressed individual serves as the US and the strength of the US may be modulated by the perception and interpretation of the mental state of the learning model and the cortical representation of empathetic pain through input from ACC and the insular cortex (Olson & Phelps, 2004, 2007). For instructional learning, while direct amygdala involvement is unlikely, it is indirectly modulated through communications of the related networks and the CS-US association.

It has also been shown at the neural, psychophysiological, and behavioral level that instructional and observational learning lead to selective processing of associated cues such as threatening compared to safe stimuli, reward and loss, or the processing of pain (Bublatzky & Schupp, 2012; Bellebaum et al., 2010; Mertens et al., 2018; Colloca, & Benedetti, 2009). For instance, enhanced feedback-related negativity and P300 were found both when observing other individuals' negative feedback as well as when receiving negative feedback in a monetary reward task (Bellebaum et al., 2010; Rak et al., 2013)

Socially induced threat elicits ERP correlates of fear processing (frontal positivity, P300, early posterior negativity, late positive potential [LPP]) and thereby modulates perceptual processing as well as behavioral outcomes of, for instance, decision-making, memory performance and attention (e.g., Baas et al., 2002; Bublatzky et al., 2010; Bublatzky & Schupp, 2012; Bublatzky et al., 2017, Bublatzky, Kavcıoğlu et al., 2020; Paret & Bublatzky, 2020; Weymar et al. 2013, 2014; Ventura-Bort et al., 2016b; Robinson et al., 2013; Deltomme, et al., 2017). The effectivity and stability of socially transmitted aversive apprehension in modulating various psychophysiological response systems has been consistently demonstrated (Bublatzky, Gerdes, & Alpers, 2014; Bublatzky et al., 2022). Defensive behavior is evident in increased activity of the somatic and autonomic nervous system for threat cues compared to safety cues, even over the course of repeated test days (e.g. enhanced skin conductance responses [SCR], cardiac deceleration, and startle reflex potentiation; Bradley et al., 2005; Bradley, Silakowski, & Lang, 2008; Bublatzky, Gerdes, White et al., 2014; Bublatzky et al., 2013; Bublatzky et al., 2018; Bublatzky et al., 2019; Costa et al., 2015; Mertens & De Houwer, 2016).

Thus, it is well established that when being in a state of aversive apprehension, defensive response programs are activated without having to experience actual threat. This threat anticipation elicits a potentiated startle reflex, heightened skin conductance as well as the activation of neural regions mediating defense behavior as a reflection of heightened perceptual vigilance (Grillon et al., 1991; Bradley et al., 2018). These measures were used as indicators for the activation of a defensive state in Study 3 and the anticipation (but not experience) of a threat. Therefore, the CS was never actually paired with the US (Costa et al., 2015; Raes et al., 2014).

Social learning can play an important role in the development and maintenance of anxiety and stress-related disorders, e.g., post-traumatic stress disorder (PTSD; Debiec & Olsson, 2017). Here, for instance, particular attention is paid to the acquisition of traumatic memories. In addition to one's own direct experience of trauma, vicarious experience (e.g., through observation) is also a trigger to be considered (see Diagnostic and Statistical Manual of Mental Disorders, DSM-5; vicariously acquired post-traumatic stress disorder, American Psychiatric Association, 2013). Social learning can be equally significant in the treatment of PTSD (Askew & Field, 2008), for instance, in situations when direct confrontation with the anxiety-provoking stimulus/situation is not possible (e.g., through exposure). Thus, no direct learning can occur that a previously threatening stimulus/situation is now safe and there is no longer any danger. Social learning would provide the opportunity of experiencing safety or the absence of threat by verbally interacting with or observing others (e.g., the therapist), which would allow for a more flexible, tailorable, and adaptable therapy design. Therefore, social learning can be a tool for inducing arousal and for establishing a threatening context in which (mal)adaptive memory processing can be investigated.

1.4 Face Perception and Recognition

Human face perception is likely the most developed visual perceptual skill in humans (Haxby et al., 2000). Humans are evolutionary primed to respond to facial stimuli automatically and can detect human faces rapidly and efficiently (Crouzet et al., 2010). How humans process faces is different to how they process non-facial objects. There are different neural systems specialized in the detection and identification of facial stimuli (Haxby et al., 2000). There is a distributed neural system in the brain for face perception with three central regions: the occipital gyri, the lateral fusiform gyrus (also involved in face identification) and the superior temporal sulcus (involved in more general face perception; Kanwisher et al., 1997; Omer et al., 2019). These neural regions respond more strongly to faces than to objects and comprise the “face” area (Natu & O’Toole, 2011).

This efficiency in processing leads to a facilitation of perception and attention directed towards faces (Palermo, & Rhodes, 2007; Farah et al., 1998; Haxby et al., 2002). Facial features carry a lot of salient (social) information, such as identity (e.g., age, gender), emotional state and intentions (Dimberg & Öhman 1996; Ekman, & Friesen, 2003). Attending to those facial features is highly relevant in social interactions and influences behavior, perception, and physiological responses (Hamm et al., 2003; Bublatzky, Gerdes, White et al., 2014). Facial features trigger (successful) social communication and interactions, as they provide information about motivation and intention. One set of markers are directed emotional expressions and posture (e.g., whether a face is pointed towards the receiver or not, Bublatzky et al., 2017). The ability to process this information might be more relevant than the skill to recognize identity (at least when the identity is unknown). There are invariant aspects of a face that specify identity and changeable aspects that facilitate social communication. Humans are very good at recognizing familiar faces even under drastically changing environmental conditions as revealed by elevated face categorization performance and differential priming effects for familiar compared to unfamiliar faces (Johnston, & Edmonds, 2009; Bruce 1982; Hill et al 1997; Burton et al 1999; Tong and Nakayama 1999; Bruce et al 2001; Sinha et al 2006, Ramon et al., 2011; Ellis et al 1990).

In contrast, the ability to recognize unfamiliar faces is surprisingly bad, and easily disrupted by changing and increasingly worsened circumstances (e.g., low lighting, low-quality pictures, variable viewpoints of the face [frontal, from the side]; Burton & Jenkins, 2011; Johnston & Edmonds, 2009; Megreya & Burton, 2006). Familiar faces activate brain regions associated with information about that person (semantic, emotional, etc.), so that the network of memory representations is enriched (Natu & O’Toole, 2011). Unlike familiar faces, unfamiliar faces are stored mainly without associated information that could facilitate the retrieval processes, even after multiple presentations. For instance, enhanced late negativity (N400) and enhanced late positivity (P600) were found for familiar compared to unfamiliar faces, even in the absence of attention being directed towards the faces (Eimer, 2000).

Furthermore, emotional (familiar and unfamiliar) faces enhance perceptional, attentional and memory processes compared to neutral faces. Increased activity in the amygdala, hippocampus, extrastriate cortex, frontal and parietal cortices was linked to the recognition of negative compared to positive faces (Keightley et al., 2010). Emotional faces trigger increased ERP positivity relative to neutral faces, even in very early processing stages (120-180ms; Schindler, & Bublatzky, 2020; Schupp et al., 2004). Threatening faces are more easily processed compared to nonthreatening, neutral, or friendly faces, with threatening faces eliciting an enhanced early posterior negativity. At later stages of stimulus processing, facial threat elicits augmented late positive potentials relative to other facial expressions (Schupp et al., 2004). A threat advantage is also evident in behavioral performance, by means of faster and more accurate detection rates (Öhman et al., 2001). Less is known, however, about a possible spillover effect from a threatening (i.e., arousing) context to an otherwise neutral face. Is an unknown face with a neutral expression better remembered when previously encountered in threatening circumstances? Preferential processing of neutral faces in a threatening context has been found for both familiar (even beloved) and unfamiliar faces. This shows enhanced activation of defensive systems (startle EMG, EDA) regardless of the familiarity status, with a robust effect, even over the course of several test days (Bublatzky et al., 2022). One of the key questions of this thesis is whether this processing benefit of (un)familiar faces also translates into enhanced memory and recognition.

1.5 Adverse Childhood Experiences

Stress- and anxiety disorders, including PTSD, which is a reaction to a traumatic event, are among the most common mental disorders (Merikangas et al., 2010). In Germany, approximately 25% of the population experience at least one such event in their lifetime (Hauffa et al., 2011), of which approximately 5 to 12% develop a trauma-related disorder, e.g. PTSD (Javidi & Yadollahie, 2012). In highly traumatized populations, such as victims of sexual abuse, the prevalence of mental disorders is higher (e.g., Spataro et al., 2004). Many individuals develop PTSD due to childhood abuse in their own homes, kindergartens, schools, and churches by close caregivers. About one in seven adults in Germany experienced sexual abuse in childhood or adolescence (Jud et al., 2017). This is associated with health, social, financial, and occupational impairments, and causes increased demands on health resources, reduced work performance, and, above all, a severely impaired quality of life (Boscarino, 2006; Giesinger et al., 2020). Therefore, a deeper understanding of the factors that influence the acquisition, maintenance, and successful treatment of trauma related disorders, such as PTSD, and trauma-related impairments, such as impaired cognitive functioning, is of high societal relevance (Solomon & Davidson, 1997).

Trauma can be self- or vicariously experienced, for example, as a witness or through a relative or close friend having been exposed to a traumatic event. Traumatic events (such as emotional and physical abuse and neglect and sexual abuse) can occur in early childhood and continue to have an impact into adulthood. A reaction to severe exposure can then lead to intense anxiety and avoidance of stimuli and situations related to the trauma. Symptoms include reliving the traumatic event (nightmares, flashbacks, intrusive thoughts), restriction of reactivity, and symptoms of increased arousal such as difficulty falling asleep/sleeping, difficulty concentrating (Bisson et al., 2013; DSM-5, American Psychiatric Association [APA], 2013). Not everyone who experiences or is indirectly affected by a traumatic event develops a trauma-related disorder. It has been suggested that psychopathology develops when a traumatic experience is processed in such a way that individuals continue to perceive a serious current and ongoing threat even in safe situations (Ehlers & Clark, 2000). This is elicited because the trauma is continually negatively appraised and autobiographical memory for that event is impaired because it has been poorly elaborated and stored without contextual information, creating false associations with threatening contexts (Ehlers & Clark, 2000). Deficient memory processes thus represent a central feature of PTSD (Brewin, 2011).

Individuals with a history of traumatic experiences report impairments in concentration and memory not only in relation to the traumatic event(s) but also in everyday life. (Distorted) Memory functions have an important role in the pathogenesis, symptomatology, and treatment of mental disorders (Vasterling et al., 1998; Johnsen & Asbjørnsen, 2008; Qureshi et al., 2011). For example, the ability to learn and retrieve neutral stimuli, especially in the verbal but also visual domain, seems to be impaired in traumatized individuals (Brewin, 2011). Furthermore, several studies have shown verbal declarative memory deficits across a variety of different trauma experiences such as war, rape, childhood abuse (Brewin et al., 2007; Johnsen & Asbjornsen, 2008). In terms of contextual factors, severely traumatized individuals show poorer recognition performance of contextual information, e.g., the background in which a stimulus is presented (Kensinger, 2007; Brewin, 2011). However, source memory has rarely been explicitly studied in the context of traumatized populations (see also Tapia et al., 2012; Brewin et al., 2012).

Deficits in working memory have been found across a wide range of tasks and within different trauma categories (Jelinek et al., 2008; Koso & Hansen 2006; LaGarde et al., 2010; Stein et al., 2002; Vasterling et al., 2002). The underlying question of this thesis is whether socially induced threat and safety contexts (via observational and instructional learning) promote or inhibit these memory systems (source and working memory) in individuals who have experienced traumatic events (i.e., adversity during their childhood). Chronic stressful events, such as early physical and emotional neglect and physical, emotional, and sexual abuse, can lead to HPA axis dysregulation (McEwen, 2002; Shea et al., 2005) and affect brain morphology involved in memory and learning in the long-term (e.g., hippocampus, prefrontal cortex; Bremner, et al., 2003). When stressors are extreme, chronic, and early in life (Heim & Nemeroff, 2001), they are associated with a reduction of cortical thickness in important developmental brain structures responsible for cognitive functioning (left lateral orbitofrontal cortex [OFC] and a region peaking in the right pericalcarine cortex; Bounoua et al., 2020; Herzog &

Schmahl, 2018) and associated with disturbed learning and memory functioning, attention and language (Savitz et al., 2007).

Cognitive biases towards threat-related and negative stimuli are common in individuals suffering from trauma and stress- and anxiety-related disorders (Dalgleish et al., 2003). Biases are, however, often content-specific (Beck & Clark, 1997; Pergamin-Hight et al., 2015) and bound to specific cognitive domains and tasks, with some biases specific for a certain disorder and for specific types of affective stimuli (e.g., threat-related, depression-related). For instance, anxiety seems to be associated with impaired/biased attention but not with mnemonic tasks and biases are stronger for pathology-congruent material (Williams et al., 1997; Mogg & Bradley, 2005). However, the whole picture of distorted cognitive processing after childhood traumatic experience is still not well understood. The aim of this thesis is to address and compare short- and long-term memory functioning during threat and safety in individuals with adverse childhood experiences (ACE).

1.6 Research Questions

As outlined above, an arousal-enhancing effect on memory has been shown for a variety of stimuli, e.g., affective words or scenes and in the context of social stimulus processing. This was also found for emotional facial expressions. Emotional faces are generally better remembered than neutral faces. This is modulated by the emotional valence as negative (e.g., angry) faces are remembered better and processed more readily than positive (e.g., happy) ones. However, these effects seem to depend on the cognitive task and the targeted memory system, attentional focus, and circumstantial factors (Keightley et al., 2010; Schupp et al., 2004).

The main takeaway is that remembering central and peripheral information is essential for organizing future adaptive behavior, especially in potentially harmful situations. The underlying mechanism is likely an enhancement of the strength of mental representations due to arousal and emotion and is moderated by the stress-system (Roozendaal et al., 2009; McGrath, 1976). However, there are mixed findings for a possible spill-over effect of an arousing source affecting the recognition of neutral items. If the central stimulus does not contain relevant informational cues, then they must be inferred from surrounding stimuli (i.e., context or source). Memory of the context serves as an indication of in how much detail memories are stored and eventually how successful learning has been. For example, a neutral-looking person is either encountered during an arousing state (i.e., threat) or during a calm situation (i.e., safety). Under which of these context conditions is the memory for the neutral faces enhanced and when is also the memory for the context itself enhanced (e.g., for socially threatening stimuli like cheater, Bell & Buchner, 2010)?

Study 1. Encoding and recognition of faces as a function of contextual settings

Therefore, the first study of this thesis aimed at clarifying the research question whether instructed threat compared to safety enhances person recognition and context memory (Weymar et al., 2013). This was assessed using a verbal threat-of-shock paradigm (Grillon et al., 1991) as well as behavioral measures and ERPs as indicators of threat processing and memory performance. Thirty healthy participants performed a combined item/source recognition task in which they incidentally learned person identity of neutral faces during encoding either in a threatening or safe context, indicated by colored background frames. Multinomial processing tree modeling was used to determine whether item (i.e., person identity) and source memory was enhanced for faces from the threatening context. For perceptual processing differences the old/new effect with an early (300-500 ms) enhanced frontal and late (> 500 ms) parietal positivity to a previously seen face compared to the new faces were assumed. This is thought to reflect familiarity and explicit retrieval processes, respectively (Rugg & Curran; 2007; Senkfor & Van Petten, 1998).

The old/new ERP effect also depends on the recollection of context (Wilding, 1999, Wilding et al., 1995). Different sources lead to differences in processing in late time windows (>

500 ms), primarily over tempo-parietal scalp sites (Wilding & Rugg 1997; Wilding, 2000; Weymar et al., 2013). Therefore, this study addresses whether threat enhances the old-new recognition effect for faces and explores whether additional source-related ERP effects (Wilding & Rugg, 1997; Kissler & Strehlow, 2017) would surface. Whether arousal has an enhancing, detrimental or no effect on memory and attention processes is dependent on a variety of factors, including memory modality and individual differences (Robinson et al., 2013). Indications of a modulating effect of (social) anxiety were observed even on unconscious threat processing, with amplified processing of threat in highly compared to low anxious participants. These findings in the healthy sample in Study 2 motivated the follow-up questions of Study 2.

<u>Study 2. Face processing, perception, and recognition under threat and its influence of ACE and anxiety</u>

Therefore, in Study 2, a sample of sixty-four participants who scored high on (social) anxiety and who have experienced some type of adversity during their childhood (abuse or neglect) completed a slightly modified version of the item/source memory task of Study 1 and an additional visual working memory task. Maltreatment during childhood is likey to have long-lasting consequences for physical, mental, and social health that persist into adulthood which, in turn, affect cognitive functioning, memory and processing of social and arousing stimuli and threat. For example, having a memory that is biased towards arousing and threatening information and having reduced differential processing of emotional stimuli has been suggested as a vulnerability factor and precursor for the development and maintenance of psychopathology (Stegmann et al., 2020). With this design, I addressed the question whether the beneficial effect of arousal on memory performance is diminished by increasing levels of ACE and (social) anxiety, depending on the memory modality.

In order to compare different socially transmitted induction methods of initial threat anticipation as a context condition, I used an observational fear and instructional threat-of-shock paradigm. Both observational and instructional threat learning lead to rapid acquisition of fear responses and to selective processing of threatening over safe stimuli (Olsson & Phelps, 2007; Bublatzky & Schupp, 2012).

Building on findings of Study 1, the old/new ERP effect was expected to emerge as a late positive potential (> 500 ms) for previously seen compared to new faces. This effect was assumed to rely on context conditions, with differential processing of the former threat compared to safe and new face-context associations. The aim was to boost memory performance in order to gain insight into the (conscious) association of context conditions and face recognition and thus be able to establish a direct connection to the electrocortical processing.

Study 3. Autonomic arousal to "unrecognized" threat identities?

Although memory for the faces was enhanced in Study 2, no context-based variation in recognition performance was discovered. Altogether, Studies 1 and 2 showed a (replicated) dissociation between perceptual processing and conscious face recognition both for healthy and traumatized participants in the item/source memory task. Study 3 aimed at demonstrating whether participants would also show autonomic nervous system activation as indicators of defensive behavior to person identities that were met in a threatening context but cannot be remembered independent of (social) anxiety levels. For this purpose, fifty participants performed a slightly modified version of the item/source memory task while electrodermal activity (EDA), and startle electromyography (EMG) were recorded. Following a transdiagnostic and dimensional approach, both healthy participants and patients suffering from (social) anxiety were assessed. I assumed enhanced autonomic arousal (threat potentiated startle reflex and skin conductance response [SCR]) to faces encountered for the second time, after first being presented in a threatening compared to a safe context, to be independent of recognition performance and individual characteristics and therefore a sign of an unconsciously primed defense mechanism (Carretié et al., 2009; Carretié, 2014).

Chapter II

Study 1: Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study.

An adapted version of this chapter has been published as “Schellhaas, S., Arnold, N., Schmahl, C., & Bublatzky, B. (2020). Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study. *Neurobiology of Learning and Memory, 174,* 107280. https://doi.org/10.1016/j.nlm.2020.107280.”

2.1 Abstract

Often the source of information is as important as the information itself. The present study examined the impact of contextual threat and safety signals (source information) on memory encoding and recognition of faces (item information). In two experimental sessions, 30 participants viewed neutral face pictures. In the encoding session, 60 faces were presented with contextual background colors (blue or green, 30 pictures each) which were verbally instructed to signal either threat-of-shock or safety. In the recognition session, the 60 old faces intermixed with 30 new faces were shown while a combined old/new recognition and source memory task was performed. During the encoding session, face processing varied as a function of contextual source information. Confirming successful threat-of-shock manipulation, threatening compared to safe face–context compounds revealed differential neural processing (early parieto-occipital and late fronto-central negativity) as well as pronounced threat ratings. During the recognition session, participants had serious problems identifying old from new faces with poor source memory. Intriguingly, however, brain activity differentiated previously seen faces from newly presented pictures (old/new ERP effect). Moreover, old faces presented within a threat context were associated with distributed late negativities compared to old safe faces. Thus, threat effects

not only emerged during face encoding (incidental learning) but also during face recognition, although no valid judgements could be made regarding the threatening or safe sources. These findings support the notion that contextual source information critically modulates person perception and recognition as a form of an expectation - based remembering in the absence of conscious recognition.

2.2 Introduction

Beneficial effects of emotional arousal on memory performance are well documented. Key findings demonstrate improved encoding and retrieval of emotionally arousing compared to neutral events (Bradley et al., 1992; Cahill & McGaugh, 1995; Dolan, 2002; Talmi, 2013). Vice versa, memory for non-emotional events is enhanced when those were encoded within an emotional context (e.g., Weymar et al., 2013; Ventura-Bort et al., 2016a; Ventura-Bort et al., 2016b). Thus, arousal associated with both the stimulus and/or the contexts can have strong effects on memory performance (Schwabe et al., 2012; Quaedflieg & Schwabe, 2018). The present study transferred these findings to social situations and person identity recognition. For instance, to remember a potentially harmful person is certainly helpful and adaptive as future negative encounters can be avoided (Bell & Buchner, 2010; Nairne et al., 2008). Here, contextual settings – such as the situation in which we met a person – can provide additional information; therefore, the modulatory effect of contextually signaled threat on person recognition is investigated.

Emotional stimuli are processed with priority and transmit important information for adaptive behavior modification. Building upon this, a strong line of research demonstrates improved memory retrieval for emotional over neutral stimuli (e.g., using emotional words, pictures or scenes; Dolcos et al., 2012; Mather & Nesmith, 2008; Kensinger & Corkin, 2003; Kensinger et al., 2006; Mather & Sutherland, 2011). These memory effects can be very robust, lasting for weeks or even years especially for aversive stimuli with implications for trauma and stress-related disorders, like for example a memory bias for anxiety-related information (McNally et al., 1989; McTeague et al., 2010; Jaworek et al., 2014). Here, event-related brain potential (ERP) studies provide important insights into the temporal dynamics and interactions of perceptual and memory processes. For instance, the recollection of previously learned (old) stimuli has been associated with an early mid-frontal (300 – 400ms) and a late parietal (> 500ms) positivity compared to new stimuli (old/new-recognition effect; Rugg & Curran; 2007; Weymar et al., 2013, 2014; Wilding & Rugg 1997; Wilding, 2000; Senkfor & Van Petten, 1998). Moreover, early frontal processing differences relate to familiarity with previously learned items and a late parietal positivity presumably reflects the explicit recollection-based remembering of item- and context-information (Weymar & Hamm, 2013; Rugg & Curran, 2007). Interestingly, the latter old/new difference varies as a function of emotional arousal and mirrors behavioral recognition performance (Barnacle et al., 2018; Weymar et al., 2014; Newsome et al., 2012; Weymar et al., 2011). Thus, contextual arousal can carry over to otherwise neutral stimuli leading to improved recognition performance (Smith et al., 2004; Martínez-Galindo & Cansino, 2017). Whether memory for contextual settings (e.g., remembering a threatening or safe environment) also benefits from arousal-mediated recognition remains an open question.

Face and person perception is a highly developed social function in humans (Adolphs, 2002). Recent research has demonstrated that facial information – such as emotional expressions, familiarity or perceived relevance – facilitate perceptual and attentional processing (Schupp et al., 2004; Schweinberger & Neumann, 2016; Eimer & Holmes, 2007; Keightley et al., 2010; Bublatzky et al., 2017). In addition, social interaction with other people always takes place within a situational context, which has been shown to modulate face perception (for a review see Wieser & Brosch, 2012). Contextual settings can selectively guide attentional resources towards emotional (facial) information, which in turn affects the memory for person identity and situational settings (Bradley et al., 1992; LaBar & Cabeza, 2006). This link between perceptual and memory processes is particularly evident in potentially threatening environments (e.g. while anticipating aversive events; Grillon et al., 1991; Bublatzky & Schupp, 2012). For instance, memory for (emotional) words is better when those were encoded during instructed threat-of-shock relative to safety condition. Moreover, this behavioral recognition effect is associated with enhanced old/new recognition effects of the event-related brain potentials (Weymar et al., 2013, 2014; Ventura-Bort et al., 2016a; Smith et al., 2004), suggesting context-based memory facilitation. However, to what degree source memory is modulated by threat is unknown; recent findings range from enhanced remembering to threat-impaired memory performance (Chiu et al., 2013) – with very few evidence concerning the contextual details of face perception.

Therefore, the present study examined the incidental learning and subsequent person memory as a function of contextual threat or safety. During an encoding session, faces were presented with contextual background colors (blue or green) signaling either threat-of-shock or safety while no memory task was mentioned (i.e., incidental learning). In the following recognition session, all previously presented (old) and new faces were presented intermixed without background frames, and participants completed an old/new recognition and source memory task. Similar to previous research (cf. Weymar et al., 2013), we predicted that a threatening (relative to a safe) context improves the encoding and consequently the recollection of faces. Moreover, multinomial tree modeling served to test source memory effects (Bayen et al., 1996), which should be better for threatening relative to safe face–context compounds. On the neural level, we predicted differential electrocortical processing for neutral faces pictures when presented during a threatening compared to a safe contextual background (i.e., threat enhanced P2 and LPP amplitudes; Bublatzky & Schupp, 2012). Regarding the recognition session, we expected pronounced late old/new differences regardless of the encoding context similar to previous research (old/new effect; Weymar et al., 2013, 2014; Gutchess et al., 2007; Rugg &

Curran, 2007 for review). Furthermore, assuming that threat-of-shock improves stimulus encoding, differential processing patterns were expected for (old) faces that had been presented within threat- compared to safety context (Wilding & Rugg, 1997; Kissler & Strehlow, 2017).

2.3 Methods

Participants

Thirty healthy participants (20 females) between the age of 20 to 26 years (*M* = 22.53, *SD* = 1.59) were recruited from the population of Mannheim and the University of Mannheim (Germany) via personal approach and advertisement in an internet platform of the university.

Questionnaire scores were within a normal range for State-Trait-Anxiety (Trait; M = 39.30, SD = 7.51; State; M = 34.37, SD = 6.08, German version of the State-Trait-Anxiety Inventory, STAI-S/T, Laux, et al., 1981), depression (BDI II; M = 6.13, SD = 5.40, Beck Depression Inventory II, BDI II, Hautzinger et al., 2006), and social anxiety (SPIN; M = 15.27, SD = 9.71; FNEK (M = 34.53, SD = 7.62; and SIAS; M = 21.97, SD = 9.71, German versions of Social Phobia Inventory, SPIN; Stangier & Steffens, 2001; Brief Fear of Negative Evaluation Scale, FNEK, Vormbrock & Neuser, 1983; Social Interaction Anxiety Scale, SIAS; Stangier et al., 1999). For the encoding session, ERP data from two participants (both female) had to be excluded because of excessive noise, one of them was also excluded from analyses of the recognition session. Thus, the final sample was N = 30 for rating and behavioral data and N = 28/29 for ERP data. Participants could choose to receive either a monetary compensation for participation (20€) or course credits. They provided written informed consent to the study protocol, which was approved by the local ethics committee.

Materials and design

Face pictures of 90 actors (45 females) were selected from the Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998) and the Radboud database (Langner et al., 2010)[1]. To focus on face identity processing, only neutral facial expressions were used. The first part of the experiment served as an encoding session (see Figure 1) in which a random selection of 60 pictures was presented (442 x 606 pixels). Half of the pictures were surrounded by a green and the other half by a blue background frame (1280 x 1024 pixels; RGB-values: 0,255,0 and 0,0,255). Each picture was presented once for 1s in random order, separated by an inter-trial interval (ITI) of in average 680ms (SD = 44.85, ranging between 620-1150ms) in which a fixation dot was presented.

In the second part of the experiment (recognition session), all 90 face pictures were presented without colored background frames in a random order (i.e., the 60 old faces from the encoding session, and 30 new faces). The participants had to decide whether a face had been presented previously with one of the backgrounds and if so, with which background (combined old/new item and source recognition task), or whether a face was new. To this end, pictures remained on the screen until the participant responded by pressing one of three keyboard buttons. The behavioral options were either a blue or a green button (corresponding to the background), or a red button for newly presented pictures. Each choice was followed by an ITI (1s, black rectangle) replacing the pictures and followed by the next face. All stimuli were presented on a 22-inch computer screen placed approximately 1m in front of the participants using the open-source software OpenSesame (Mathôt et al., 2012).

---

[1] af01nes.jpg, af02nes.jpg, af04nes.jpg, af05nes.jpg, af06nes.jpg, af07nes.jpg, af08nes.jpg, af09nes.jpg, af11nes.jpg, af13nes.jpg, af14nes.jpg, af15nes.jpg, af17nes.jpg, af18nes.jpg, af19nes.jpg, af20nes.jpg, af22nes.jpg, af23nes.jpg, af24nes.jpg, af25nes.jpg, af26nes.jpg, af27nes.jpg, af28nes.jpg, af29nes.jpg, af30nes.jpg, af31nes.jpg, af32nes.jpg, af34nes.jpg, af35nes.jpg, am17nes.jpg, am18nes.jpg, am21nes.jpg, am22nes.jpg, am23nes.jpg, am25nes.jpg, am26nes.jpg, am28nes.jpg, am29nes.jpg, am30nes.jpg, am31nes.jpg, am32nes.jpg, am34nes.jpg, am35nes.jpg, bm12nes.jpg, bm16nes.jpg, bm24nes.jpg, am01nes.jpg, am02nes.jpg, am03nes.jpg, m04nes.jpg, am05nes.jpg, am06nes.jpg, am07nes.jpg, am08nes.jpg, am09nes.jpg, am10nes.jpg, am11nes.jpg, am13nes.jpg, am14nes.jpg, Rafd090_01.jpg, Rafd090_02.jpg, Rafd090_08.jpg, Rafd090_12.jpg, Rafd090_14.jpg, Rafd090_16.jpg, Rafd090_18.jpg, Rafd090_19.jpg, Rafd090_22.jpg, Rafd090_26.jpg, Rafd090_31.jpg, Rafd090_32.jpg, Rafd090_37.jpg, Rafd090_57.jpg, Rafd090_58.jpg, Rafd090_61.jpg, Rafd090_05.jpg, Rafd090_07.jpg, Rafd090_09.jpg, Rafd090_10.jpg, Rafd090_15.jpg, Rafd090_23.jpg, Rafd090_24.jpg, Rafd090_25.jpg, Rafd090_28.jpg Rafd090_30.jpg, Rafd090_33.jpg, Rafd090_36.jpg, Rafd090_38.jpg, Rafd090_46.jpg, Rafd090_71.jpg

Procedure

After the EEG sensor net was attached, a fake stimulation electrode was placed at the left inner forearm. Thus, similar to previous studies (e.g., Costa et al., 2015), the mere attachment of an electrode – without actually experiencing any electric stimulation – was used to induce aversive anticipations. The participants were verbally instructed that the intensity of the electric stimulation, given during the experiment, had been rated as ‘maximal unpleasant but not painful’ by participants in former studies. Participants were then instructed that they might receive a maximum of three electric shocks whenever a specific background color was presented. Thus, one colored background frame served as a threat cue (e.g., green indicated the possibility of receiving shocks) whereas the other color signaled safety (e.g., blue indicated the absence of any shocks). Color assignment to conditions (threat, safety) was counterbalanced across participants.

To examine the anticipation (but not experience) of threat, no shocks were administered during the experiment. Previous research has consistently shown the effectivity and stability of such threat-of-shock instructions in modulating various psychophysiological response systems. For instance, enhanced activity of the somatic and autonomic nervous system has been shown for instructed threat relative to safety cues (e.g., enhanced skin conductance responses, heart rate deceleration, and potentiated startle responses; Bradley, Moulder, & Lang, 2005; Bublatzky, Gerdes, White et al., 2014; Bublatzky et al., 2013; Bublatzky et al., 2018; Bublatzky et al., 2019; Costa et al., 2015). Moreover, instructed threat modulates perceptual processing and behavioral responding such as decision-making and memory performance (e.g., Baas et al., 2002; Bublatzky et al., 2010; Bublatzky & Schupp, 2012; Bublatzky et al., 2017, Bublatzky, Kavcıoğlu et al., 2020; Paret & Bublatzky, 2020; Weymar et al. 2013; Ventura-Bort et al., 2016b).

For the encoding session, participants were instructed to attentively watch all pictures presented at the screen, without excessive moving or blinking. Focusing on incidental learning, no recognition test was mentioned. After completing the encoding session, the shock electrode was removed and the participants rated the two background colors (without the face stimuli) with regard to valence and arousal using the self-assessment manikin (SAM; Bradley & Lang, 1994; ranging from 1 to 9) and perceived threat (ranging from 0 to 10 on a Likert-Scale). Because previous research showed the comparability of the background colors (green and blue) as threat/safety signals, no additional baseline rating was completed before the encoding session (Bublatzky et al., 2013, 2014, 2020). While separating the encoding and recognition session, this break also served to assess questionnaire data regarding (social) anxiety and depression.

In the following recognition session, participants had to decide whether a picture was presented with a green or blue background or was new. If participants were sure about the previous occurrence of a picture but unsure about the background color, they were instructed to guess the background color. There was no time limit for the response and no feedback was given regarding the accuracy of their choice. At the end of the experiment, a brief post-interview was performed and participants were debriefed.

Data recording and reduction

Electrocortical activity was recorded using a 64-channel actiCap system (BrainProducts, Munich, Germany). Ag/AgCl active electrodes were mounted into an elastic cap using a 10-10 electrode placement standard with FCz as reference electrode (Falk Minow Services, Herrsching, Germany). The EEG was recorded continuously with a sampling rate of 500 Hz and filtered online from 0.1-100 Hz using Vision Recorder acquisition software and BrainAmp DC amplifiers (BrainProducts). The impedance of all electrodes was kept below 10 kΩ. Offline data analyses were done using Brain Vision Analyzer 2.0 (BrainProducts) and EMEGS (Peyk et al., 2011), including conversion to an average reference, low-pass filtering at 30 Hz, artifact detection, sensor interpolation, and baseline correction as described by Junghöfer et al. (2000). Stimulus-synchronized epochs were extracted and lasted from 100ms before to 1000ms after stimulus onset in the encoding and recognition session. Finally, separate average waveforms were calculated for the experimental conditions Instruction (threat, safety) and Recognition (old, old plus source, new) for each sensor and participant.

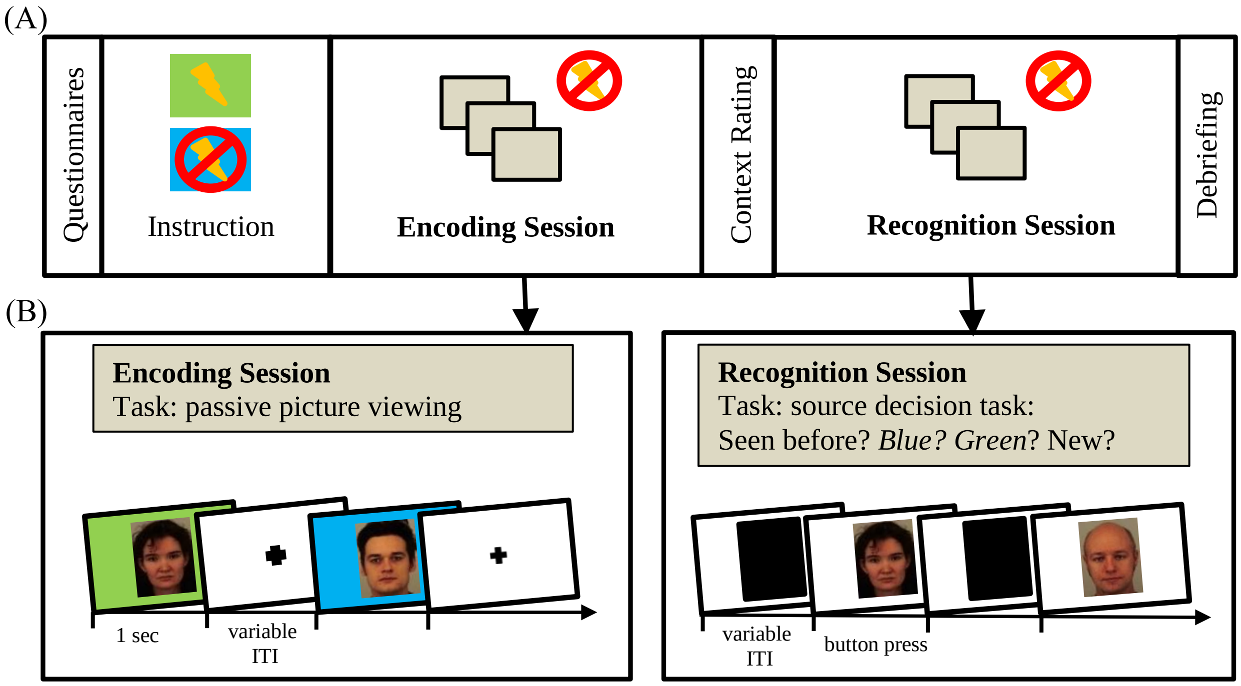


Figure 1. Schematic illustration of the experimental procedure (A) and stimulus presentation and experimental task (B). During the encoding session, 60 pictures of male and female faces displaying neutral expressions were presented for 1 s each (variable ITI). Colored picture frames (blue or green) served as verbally instructed threat-of-shock or safety cues. Participants' task was to attend to all pictures, the following recognition task was not mentioned (incidental learning). During the recognition sessions, the 60 old and 30 new faces were presented intermixed without background frames. Participants' task was to decide whether the face was shown before and with which frame (combined old/new item and source recognition task), or whether it was not previously presented.

Data analysis

Self-report data. Separate *t*-tests were calculated for valence, arousal, and threat ratings in order to confirm expected differences between the threat-of-shock and safety conditions.

Behavioral data. To quantify memory performance during the recognition session, hit rates (correct responses divided by all responses in target trials) and false alarm rates (incorrect responses divided by all responses in non-target trials) were computed for old and new faces and analyzed by separate *t*-tests. Source-memory for the encoding context of faces was analyzed by conducting *t*-tests using the conditional source identification measure (CSIM), which is the number of all correct source identifications divided by the overall number of faces that were identified as "old" (Bell & Buchner, 2011; Bröder & Meiser, 2007). For example, the CSIM of a face from a threatening context is calculated by dividing the number of correct source allocations for threat (i.e., the response 'threat' given a face stemming from a threatening context) by the total number of the source allocation for threat (i.e., the response 'threat' given a face stemming from a threatening and a safety context).

To disentangle item memory, source memory and guessing biases for recognition and source judgments, multinomial processing tree (MPT) modelling was used (Batchelder & Riefer, 1999; Bayen et al., 1996).

MPT models are well-established mathematical models which use discrete observable states to draw conclusions of non-observable underlying cognitive processes. For example, in a source monitoring task, as used in the present study, participants´ observable responses (i.e., correct or incorrect old/new recognition and source identification) reflect the workings of various latent cognitive states such as knowledge or guessing (Batchelder & Riefer, 1990). MPT models assume that the observed categorical responses follow a multinomial distribution. The model parameters represent transition probabilities between latent cognitive states (Matzke et al., 2015 for detailed reviews, see Batchelder & Riefer, 1999; Erdfelder et al., 2009; Arnold, Bayen, & Smith, 2015). These assumptions can be depicted as decision trees with observable frequency data at the end of the branches and the branches itself describing the model parameters (see Figure 2).

The MPT model we used is the two-high-threshold model of source monitoring (2HTSM; Bayen et al., 1996). Items are presented with one of two sources either with Source A or with Source B, more specifically, in the present experiment either within a threatening or safe context. In the recognition phase, participants also see new items (i.e., items that were not presented with one of the sources). The model consists of three trees – each representing one item (i.e., face) type: (1) faces presented in the threat condition, (2) faces presented in the safety condition, and (3) new faces (i.e., that were not presented during the encoding phase). The answer categories (answering that the face was presented with the threat color, safety color or is new for each item type) can be reached via various latent cognitive states[2]. Participants correctly recognize a face as old or new with probability *D*. With probability d_T participants also correctly recognize that the face originated from the threatening source, with probability d_S they correctly recognize the face stemming from a safe source. If they do not remember the source of the face (probabilities 1-d_T and 1-d_S for threat and safety, respectively), they have to guess. With probability *g*, participants guess that the face originates from a threatening source and with probability 1-*g* that it derives from a safe source. If participants do not recognize that a face was old or new (probability 1-*D*), they guess with probability *b* that a face was old and therefore with probability 1-*b* that a face was new. If they do not remember whether a face is old or new, they must also guess the source. Again, with probability *g* they guess that a face was previously presented with a threatening and with probability 1-*g* with a safe source. The overall probability of a certain response category then results from the sum of the probabilities of the individual branches, for example, the probability for answering “Threat” when an item stems from a threatening source is p(“Threat”) = $D*d_T + D*(1-d_T)*g+(1-D)*b*g$.

---

[2] Please note, that the full model described by Bayen et al. (1996) is not identifiable (eight parameters and only 6 free response categories). We therefore used Submodel 5d (Bayen et al., 1996), in which some parameters are restricted to be identical. This choice was based on theoretical considerations in order to make the model identifiable and obtain unique parameter estimates. Specifically, it assumes that item memory as well as source memory are equal for both sources and that the probability of noticing that an item is new is equal to the probability of recognizing an item as old (see Figure 2). In a pilot study (Arnold et al., in preparation), we confirmed that this submodel is suitable for the present paradigm with four sources.

In MPT models, it is assumed that observations are equally distributed over response categories and relies on aggregated data (Batchelder & Riefer, 1999). However, these assumptions might be violated as a result of differences between participants and may lead to biased parameter estimates. In recent years, hierarchical Bayesian extensions have been developed to account for participant heterogeneity by assuming that the individual parameters follow a continuous hierarchical distribution (e.g., Heck et al., 2018; Klauer, 2010; Smith & Batchelder, 2010). Importantly, these approaches allow including predictors in the MPT model and estimating a relationship between model parameters and other data such as individual questionnaire scores (e.g., Arnold, Bayen, & Böhm, 2015; Schaper et al., 2019). Therefore, model parameters were estimated on group level as well as on individual level.

ERP data. ERP data was analyzed separately for the encoding and the recognition session. Based on previous research and inspection of the waveforms, statistical analyses were computed using sensors time windows in which differences between the respective conditions were maximal.

For the encoding session, statistical analyses were computed using two early time windows (160-170ms and 230-300ms) for parieto-occipital sensors (PO7 and PO8) and a later time window for fronto-central sensors (472-632ms; C1, C2, Fc1 and Fc2). Encoding-related ERP data were analyzed in a repeated measure ANOVA using the factor instruction (threat, safety).

For the old/new effect in the recognition session, ERP data were segmented and categorized according to several variables based on memory performance in reference to the encoding session. (1) Faces which were successfully recognized as previously presented, regardless of the memory of additional source details (old faces), (2) Faces successfully recognized as previously presented, additionally successful source recognition (old faces plus source), and (3) faces successfully recognized as not previously presented (new faces). This approach is in line with previous studies (Awipi & Davachi, 2008; Gardiner & Java, 1993; Tulving, 1983; Mattarozzi et al., 2015). Please note, because ERP data relies on memory performance, different trial numbers were obtained for each participant and condition. For Type (1) events, there were on average 35.17 (SD = 12.15, min/max = 11/53) trials per participant. For Type (2) events, the number of trials was further reduced (M = 18.87, SD = 6.81, min/max = 7/30). The average number of trials for Type (3) events was 16.07 (SD = 6.53, min/max = 5/28).

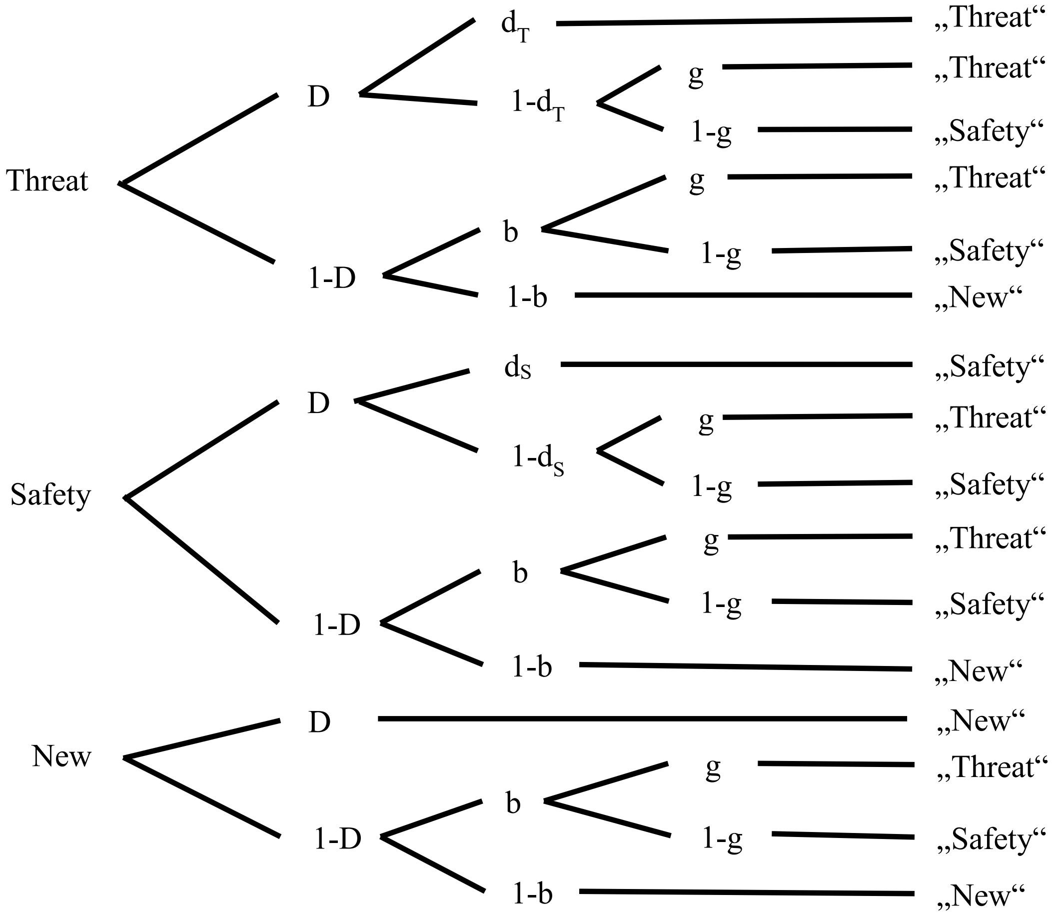


Figure 2. Submodel 5d of the two-high-threshold model of source monitoring. D probability of detecting that an item is old/new; d_T probability of correctly remembering the shock source of an item; d_S probability of correctly remembering the safe source of an item; g probability of guessing that an item is from the shock source; b probability of guessing that an item is old; adapted from Arnold et al., 2013; originally by Bayen, Murnane, & Erdfelder, 1996.

Based on previous research and waveform inspection, mean ERP amplitudes for old/new differences were analyzed in a time window of 700-760ms over parieto-occipital sensors (POz, PO3 and PO4) using only old faces and new faces trials. Additionally, mean ERP amplitudes were analyzed in a time window of 500-800ms over fronto-central sensors (Fz, and FCz) for old faces plus source and new faces trials, in order to investigate whether correct source identification influences the old/new effect.

To examine instruction effects in the recognition session, differences in the mean ERP amplitudes between new faces and those previously encoded during a threat or safe context were analyzed in a time window between 420-520ms over central sensors (Cz, C1 and C2) and in a time window of 716-900ms over parieto-occipital sensors (PO3, PO4 and Poz). Since only trials with correctly recognized faces and source were considered (old faces plus source (2), separately for the threat and safety context, the number of trials included in the analyses was rather low. For the safe condition an average of 10.37 (SD = 4.07), for the threat condition an average of 8.50 (SD = 3.83) and for the new condition an average of 16.07 (SD = 6.53) of trials were analyzed.

For the old/new effect, two repeated measure ANOVAs with the factor Recognition (old, new) were conducted for either only correctly recognized old faces (1) or old face-source compounds (2). For the instruction effect, a third repeated measure ANOVA was computed for analyzing recognition ERP data regarding anticipatory threat including the factor Instruction (threat, safety and new condition) using only old faces plus source (2) trials. For exploratory reasons, questionnaire measures were tested as covariates.

For effects involving repeated measures, the Greenhouse Geisser procedure was used to correct violations of sphericity and Bonferroni correction was used for multiple comparisons. As a measure of effect size the partial η^2 (η_p^2) is reported.

2.4 Results

Self-report data

Similar to previous research, verbal threat instructions successfully induced aversive apprehensions. As predicted the threatening context was rated more unpleasant, arousing and threatening compared to the safe context, *ts*(29) = -5.09, 5.47, and 5.90, *ps < .001 (*M_{threat} = 4.27 [SD = 1.74], 4.80 [SD = 2.01], 3.70 [SD = 2.62]; M_{safety} = 5.77 [SD = 1.38], 3.13 [SD = 1.89], 1.50 [SD = 1.87]). Moreover, a post-interview at the end of the experiment indicated that our threat-of-shock manipulation was credible over the whole period of the encoding phase. 93.3 % reported that instructed threat context made them feel more tense, nervous, expectant and attentive relative to the safety context. When asked about changes across the course of the encoding session, 16.7% reported that the anticipation of receiving an electric shock remained constant, 76.6% reported a decline, and 6.7% reported that the anticipation increased towards the end of the encoding session.

Behavioral data

The memory performance for old (item recognition and source identification) and new faces is listed in Table 1. Overall, hit rates for face recognition for both new and old faces – regardless of the encoding context – were low. Neither hit rates nor false alarm rates differed between old and new faces, *t*s(29) = -0.58 and 0.69, *p*s = 0.57 and 0.50, indicating that participants were not able to differentiate between previously presented and new faces.

Regarding source memory, the analysis of the CSIM did not reveal a significant effect, *t*(29) = 0.66, *p* = 0.51, indicating that source identification did not differ between a threatening and a safe context.

Table 1. *Recognition and source memory performance (“traditional/conventional” measures).*

	Old		New	Old
	Threat	Safe		Threat & Safe
Hit rate	0.29 (0.13)	0.35 (0.24)	0.54 (0.22)	0.59 (0.20)
False alarm rate	0.23 (0.13)	0.27 (0.12)	0.41 (0.20)	0.46 (0.22)
CSIM	0.53 (0.15)	0.56 (0.13)	N/A	N/A
CSIM both sources	0.54 (0.09)		N/A	N/A

Note. Numbers represent percentages (and SD). CSIM = conditional source identification measure.

MPT model analyses were calculated using the latent-trait approach (Klauer, 2010) as implemented in the R-package TreeBUGS (Heck et al., 2018). TreeBUGS relies on Bayesian Modeling and Marcov chain Monte Carlo (MCMC) sampling for calculating the parameter estimates. The reasons for Bayesian modeling are not only philosophical (for a discussion see Rouder & Lu, 2005) but also practicable ones, especially in nonlinear hierarchical modeling. MCMC sampling is a method for finding the posterior distribution (for details see Heck et al., 2018). The algorithm was run with 50,000 iterations using the first 10,000 iterations removed as burn-in period. For all parameters, the potential scale reduction factor was < 1.06, indicating good convergence. Group means of the parameter estimates are presented in Table 2. Questionnaire data were entered as continuous predictors (implemented in TreeBUGS by a linear regression on the probit scale) with TreeBUGS standard priors, that is, weakly informative, multivariate Cauchy priors.

As can be seen in Table 2, the probability of recognizing an item as old and new (D = 0.12) is very low (in similar paradigms D parameters of 0.5 and higher are reported, e.g., Meiser & Bröder, 2002; Cooper et al., 2016). Because of the poor performance in detecting that an item was old, consequently, the large Bayesian confidence intervals (BCIs) indicate increasing uncertainty for the estimation of the source memory parameters. The source guessing parameter is in accordance with the probabilities of the event. The proportion of items presented with a threatening context is half (g = 0.54), therefore there is no guessing bias towards a specific source. The ratio of old to new items in the recognition task is 2:1. There was a conservative guessing bias towards new items present (b = 0.52).

Table 2. *Posterior distributions of the parameters of the hierarchical distributions.*

	M [95 % BCI]	SD
D	0.12 [0.08 – 0.15]	0.02
d_T	0.32 [0.01 – 0.90]	0.25
d_S	0.53 [0.05 – 0.97]	0.27
b	0.52 [0.42 – 0.62]	0.05
g	0.54 [0.50 – 0.59]	0.02

Note. D = probability of item recognition, d_T= probability of remembering that an item was presented in a threatening context, d_S = probability of remembering that an item was presented in a safe context, b = probability of guessing that an item was old, g = probability of guessing that an item was presented in a threatening context; BCI = Bayesian confidence interval.

All BCIs of the correlations included zero indicating that there were no significant correlations between any of the parameters. However, significant correlations were observed between the parameters for guessing that a picture was old (i.e., parameter *b*) and various of the questionnaires (STAI-S: r = -.35 [-.43; -.26], BDI: r = .20 [.11; .29], SIAS: r = -.21 [-.30; -.12], SPIN: r = -.24 [-.32; -.15], FNEK: r = -.17 [-.26; -.08]. Please note, values in square brackets represent the BCI of the prediction, BCIs not including zero imply a relationship). The more (socially) anxious the participants were, the more likely they guessed that an item was new, indicating a more conservative guessing bias. For depression scores, the pattern was reversed, specifically, the higher the BDI-score the more likely the participants guessed that an item was old.

ERP data

Encoding session: Instruction effect. Differential processing of threat compared to safety faces was observed during the encoding session (Figure 3). This is revealed over parieto-occipital sensor sites showing a more pronounced negativity for threat compared to safety faces in two early time windows (160-172ms and 230-300ms), Instruction $F(1,28) = 5.35$, $p < 0.05$, $\eta_p^2 = 0.17$ and $F(1,27) = 5.74$, $p < 0.05$, $\eta_p^2 = 0.18$. Interestingly, for the earlier time window, this effect varied as a function of inter-individual differences in social anxiety as measured by SIAS and the SPIN scores, $F_{\mathrm{SIAS}}(1,27) = 4.44$, $p < 0.05$, $\eta_p^2 = .17$, and $F_{\mathrm{SPIN}}(1,27) = 8.91$, $p < 0.01$, $\eta_p^2 = .30$. Follow up analyses revealed that threat/safety differences were more pronounced, the more socially anxious the participants were.

Moreover, pronounced negativity was observed for threat compared to safety faces over fronto-central sensors in a later time window (472-632ms), Instruction $F(1,27) = 9.60$, $p < 0.01$, $\eta_p^2 = .36$. No further covariation effects were found (Fs < 1).

Recognition session: Old/new effect. Focusing on correctly recognized trials – for both old (1) and new (2) faces – significant old/new recognition effects were found despite the poor behavioral performance in conscious item recognition. Correctly classified new relative to old faces were associated with enhanced positivity over parieto-occipital sensor sites in a late time window (700-760ms), Recognition $F(1,28) = 4.38$, $p < .05$, $\eta_p^2 = .12$. This finding contradicts previous research (Kissler & Strehlow, 2017; Weymar et al., 2013; Rugg & Curran, 2007; Danker et al., 2008). However, the overall performance in the old/new recognition task was very poor. Consistent with previous approaches (Wilding et al., 1995), follow-up analyses relied only on trials in which both the faces and respective sources were correctly identified (old faces plus source; Figure 4). Here, significant old/new effects emerged over fronto-central sensors in a late time window (500-800ms), Recognition $F(1,28) = 7.43$, $p < 0.05$, $\eta_p^2 = 0.21$, with more positivity for old faces. No covariation with the questionnaire scores were found, Fs$(2,56) < 4.20$, ps > .0.05, $\eta_p^2 < .16$.

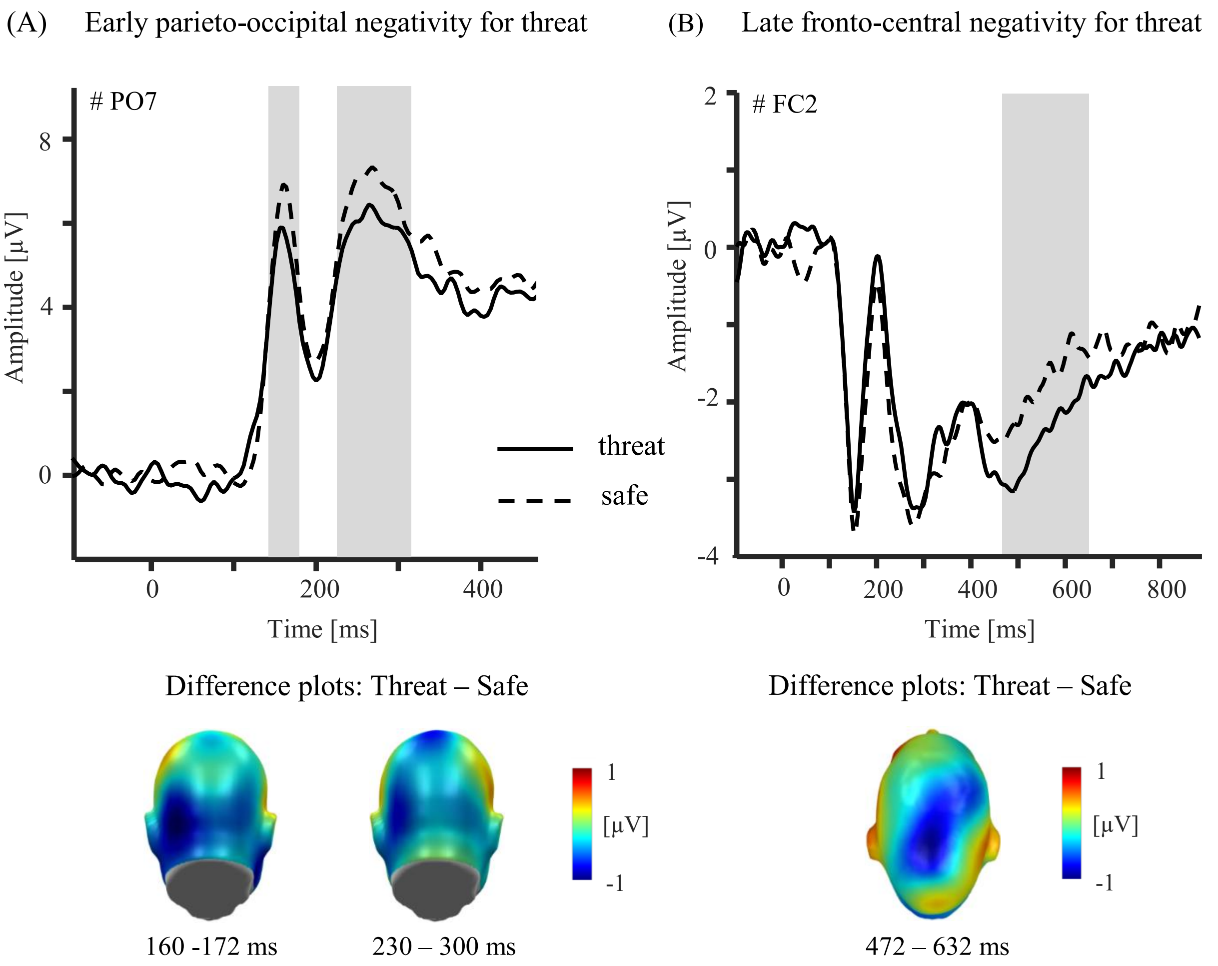
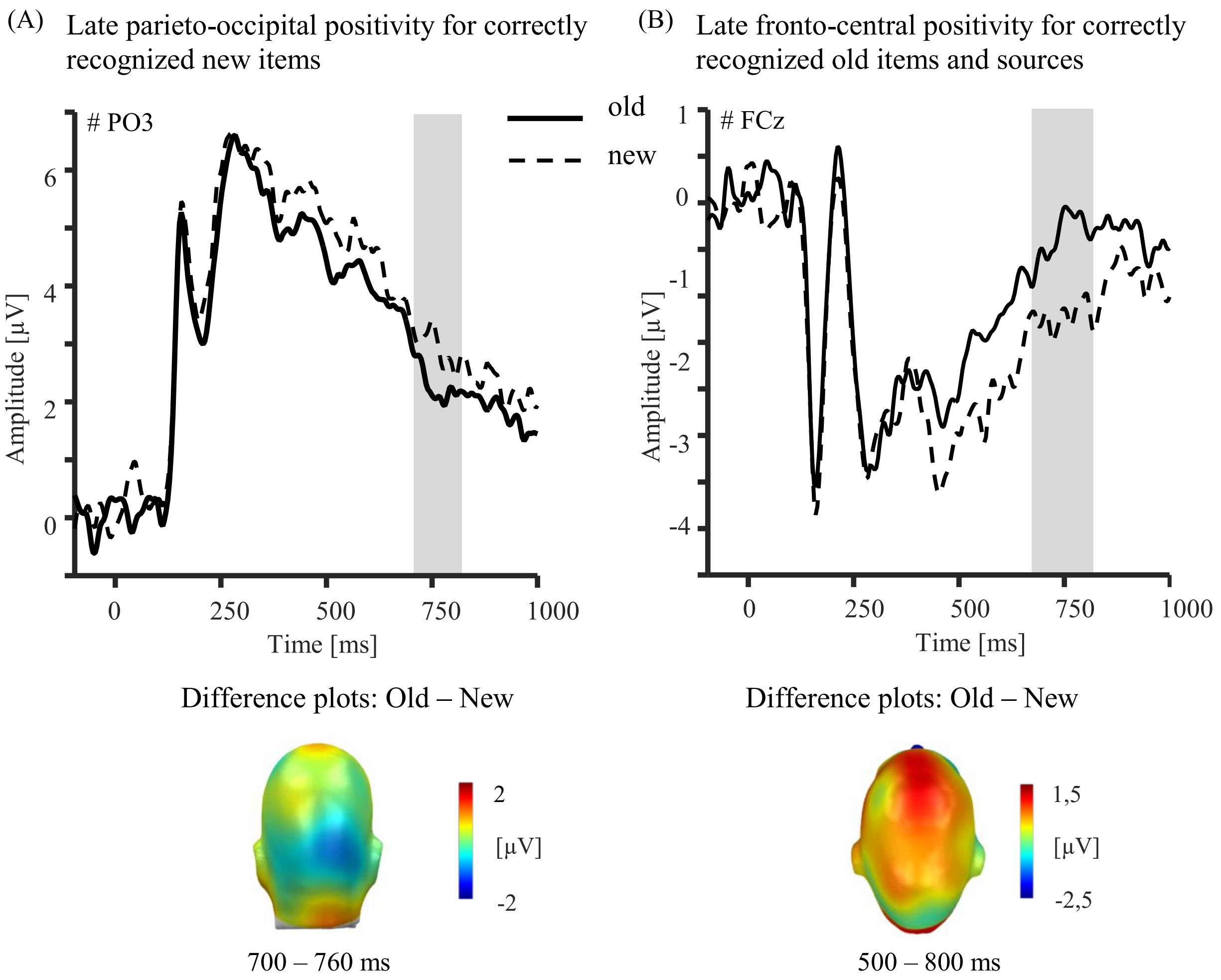


***Figure 3*. Encoding session: Instruction effect**. Illustration of the main effect Instruction in the encoding session. (A) Grand averaged ERPs prompted by faces presented with threat-of-shock or safety for an exemplary parieto-occipital sensor and topographical difference maps (threat – safe) displaying the averaged time intervals (162-175 ms and 230-300 ms) plotted on the back of a model head. (B) Grand averaged ERPs for faces presented within threat-of-shock or safety for an exemplary fronto-central sensor and a topographical difference map (threat – safe) displaying the averaged time interval (472-632 ms) plotted on a top view of a model head.

Recognition session: Instruction effect. Contextual threat effects emerged in the recognition session in a mid-latency time window (420-520ms) over central sensor sites, Instruction $F(2,56) = 3.33$, $p < .05$, $\eta_p^2 = .11$. When encoded within a threatening context, old faces were associated with enhanced negativity compared to old safety faces, $F(1,28) = 6.27$, $p < .05$, $\eta_p^2 = .18$. No significant differences were observed for old (threat or safety) compared to new faces, $Fs(1,28) < 2.72$, $ps > .11$, $\eta_p^2 = .09$ (Figure 5). Similar to the encoding session, this context effect covaried with social anxiety scores (SIAS), $F(2,56) = 3.30$, $p < .05$, $\eta_p^2 = .13$, showing more pronounced threat-new differences in more socially anxious participants.

Furthermore, an additional overall instruction effect emerged in a late time window (716-900ms) over parieto-occipital sensor sites, Instruction $F(2,56) = 3.26$, $p < .05$, $\eta_p^2 = .10$, with post-hoc tests indicating a pronounced negativity for threat-faces compared to newly presented faces, $F(1,28) = 6.40$, $p < .05$, $\eta_p^2 = .19$. No significant differences for threat and new faces compared to safe faces were observed, $F(1,28) = 3.01$, $p = .09$, $\eta_p^2 = .10$, and $F < 1$, respectively.


(A) Late parieto-occipital positivity for correctly recognized new items

(B) Late fronto-central positivity for correctly recognized old items and sources

Difference plots: Old – New

700 – 760 ms

500 – 800 ms

***Figure 4*. Recognition session: Old/New effect**. Illustration of the old/new effect in the recognition session. (A) Grand averaged ERPs prompted by old faces, which were recognized correctly, regardless of source memory, and new faces for an exemplary parieto-occipital sensor and a topographical difference map (old - new) displaying the averaged time interval (700 - 760 ms) plotted on a back of a model head. (B) Grand averaged ERPs for correctly classified old faces with correct source allocation, and new faces for an exemplary fronto-central sensor and a topographical difference map (old – new) displaying the averaged time interval (500 – 800 ms) plotted on a top view of a model head.

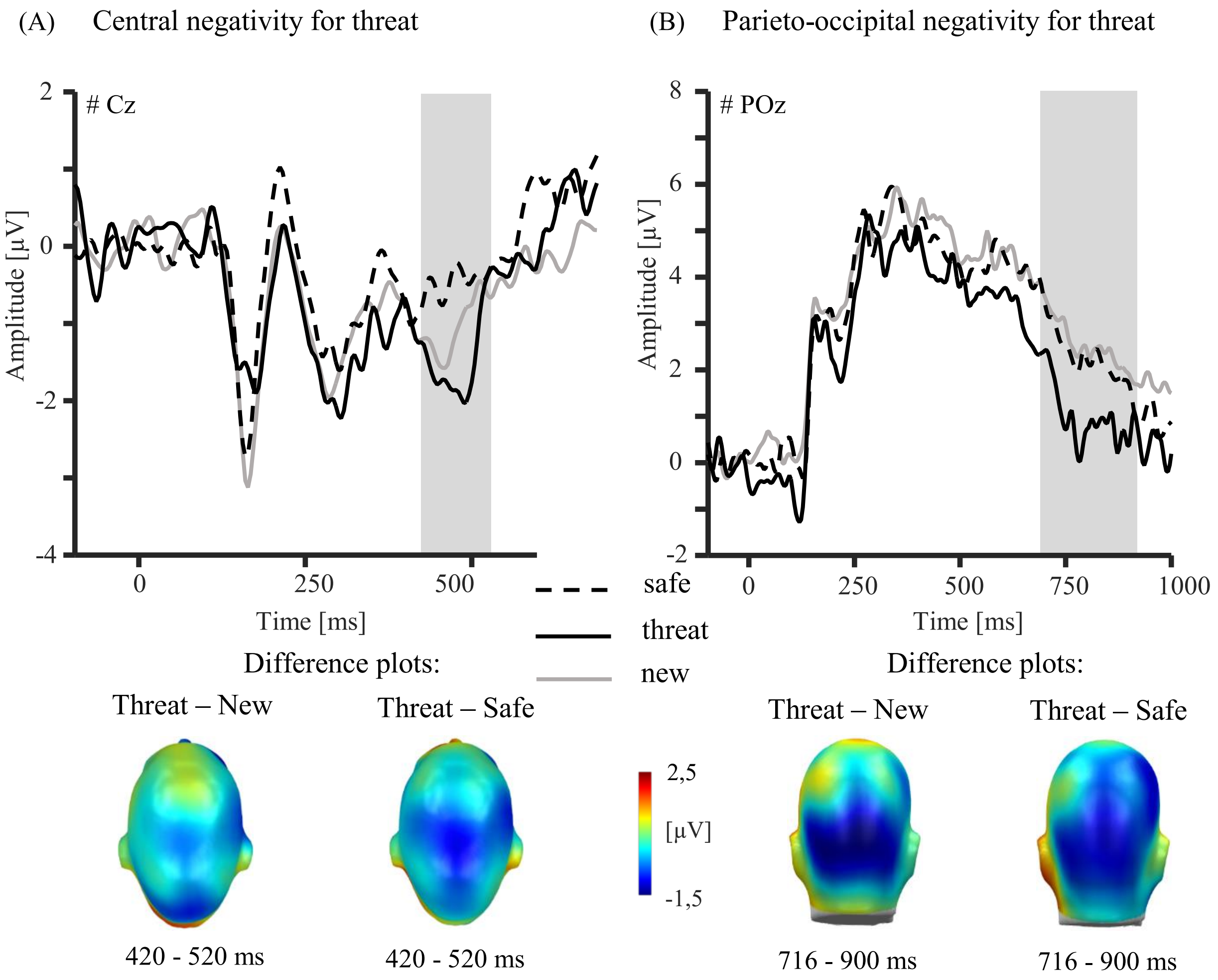


***Figure 5*. Recognition session: Instruction effect**. Illustration of the main effect of instruction in the recognition session. (A) Grand averaged ERPs for faces which have been previously presented within threat-of-shock or safety in the encoding session and new faces for an exemplary central sensor and topographical difference maps (threat – new; threat – safe) displaying the averaged time interval (420 - 520 ms) plotted on top views on a model head. (B) Grand averaged ERPs prompted by faces previously presented with threat-of-shock or safety and new faces for an exemplary parieto-occipital sensor and topographical difference maps (threat – new, threat - safe) displaying the averaged time interval (716 - 900 ms) plotted on a back view of a model head.

2.5 Discussion

Meeting a person in a threatening context can later modulate the recognition of this person (i.e., item memory). Moreover, little is known about the identification of the context in which we previously met a person (i.e., source memory). Here we examined recognition performance and electrocortical processing of multiple face identities who have been encoded either in a threatening or safe context. Differential neural processing and rating data indicate successful implementation of threat and safety context conditions during face encoding. Although memory for face identities was poor during a later recognition session, brain activity differentiated old from new faces and showed source selective face processing. Together, these findings support the notion that contextual settings critically modulate person perception even in the absence of conscious recognition.

Encoding phase with incidental learning. When passively viewing neutral faces – without an instruction to memorize face stimuli (incidental learning) – aversive apprehensions modulated face processing during the encoding session. As predicted, self-reported ratings confirm our threat-of-shock manipulation and show that the threat context was perceived as more threatening, arousing and unpleasant relative to the safety condition. Regarding electrocortical processing, early occipito-temporal negativities (160-172ms and 230-300ms) differentiated faces that were presented within a threatening relative to a safety context. These findings are in line with previous research showing early negative-going components (N170, EPN) as sensitive to threatening facial information (e.g., facial expressions of anger; Hinojosa et al., 2015; Schupp et al., 2004). Moreover, other studies observed modulated EPN amplitudes also as a function of emotionally relevant contextual information. For instance, pronounced EPN were observed when anticipating social events such as meeting a person or giving a speech (Bublatzky et al., 2014; Wieser et al., 2010) or within social group situations (Bublatzky et al., 2017). These findings have been suggested to reflect early attentional tagging of motivationally relevant information for detailed subsequent processing (Schupp et al., 2003; Schupp et al., 2004).

Regarding later stimulus processing, however, the observation of pronounced negativities over fronto-central regions (472-632ms) during the encoding session was a rather unexpected finding. Whereas most research has demonstrated enhanced late positive potentials to threatening information (e.g., Bublatzky & Schupp, 2012; Schupp et al., 2004), several studies from the cognitive domain have observed sustained negative-going waveforms. For instance, in the field of attention, memory and language processing various negative-going waveforms were considered to reflect facilitated stimulus processing (e.g., mismatch negativity, N400; Kutas & Federmeier, 2000; Näätänen, 1995; Potts & Tucker, 2001; N2pc, stimulus-preceding negativity, SPN; Kausche & Schwabe, 2020). Moreover, one recent study observed a sustained negative potential (occipital; 80-580 ms) for pleasant picture materials mismatching the concurrently presented threatening background (Bublatzky et al., 2010). Taken together, these effects indicate that instructed threat modulates both early perceptional processes as well as more elaborate processing of otherwise neutral social stimuli, during the mere passive viewing and encoding (incidental learning) of unknown faces (Baas, et al., 2002; Böcker et al., 2004; Bublatzky & Schupp, 2012; Weymar et al., 2013).

Face recognition and source-memory. For the recognition session, a combined old/new-recognition and source memory task was performed. Participants were asked to identify within which background (source) they had seen the face identities previously or whether they had not seen them before. Using multinomial processing tree (MPT) modelling, behavioral item- and source-recognition as well as guessing parameters were estimated. The old/new recognition parameter revealed that participants were not able to discriminate between old and new faces. Consequently, there was not enough statistical power to estimate meaningful source-memory parameters as a function of contextual threat or safety (Bisby & Burgess, 2014). Despite the lack of behavioral threat-effects on item or source recognition, however, an old/new processing difference emerged for electrocortical brain activity. In a first set of analyses (common strategy; Weymar et al., 2013, 2014; Rugg & Curran, 2007), all correctly classified old and new faces were taken into account, regardless of the accuracy of the source identification. Contradicting previous research, new faces were associated with an enhanced late (700-760ms) parieto-occipital positivity compared to the old faces. However, accounting also for correct source judgements (cf. Wilding & Rugg, 1996; Wilding, 1999; Doerksen & Shimamura, 2001), a second set of analyses was conducted. For trials in which both item- and source-information were correctly classified (i.e., face identity and threat/safety context), an enhanced late positivity (500-800ms) was found over fronto-central brain regions for previously encountered faces. Thus, late positive potentials towards previously seen faces may reflect a processing advantage for old item and source information during recognition (Wilding, 1999; Senkfor & Van Petten, 1998; Mecklinger, 2006).

These findings may be interpreted from the perspective of a dual-process theory as proposed by Rugg and Curran (2007). According to this model, two memory processes characterize old/new recognition effects as a function of face familiarity (indicated by early frontal ERP components) and second recollection processes (reflected by late parietal ERP correlates; for an overview see Rugg & Curran, 2007). Whereas familiarity-based recognition is considered non-contextual (Wilding & Rugg, 1996), face recollection increases with the number and quality of additional information from the encoding context (e.g., retrieval of contextual cues; Addante et al., 2012). The present ERP data might indicate that participants actively recollected the faces and thus person identity rather than merely having a sense of familiarity. However, this needs to be considered with caution, because the behavioral recognition performance did neither indicate old/new nor source memory effects. Likely, the fast presentation of face–context compounds (1s with a brief ITI of ~680ms) and incidental learning tasks (no memory task was mentioned) prevented any better recognition performance.

Nevertheless, electrophysiological correlates of source classification – even without conscious recognition – support a recollection-based account. Assuming that contextual information enhances perceived relevance to an observer (e.g., explicit instructions or accompanying faces; Bublatzky & Schupp, 2012; Bublatzky et al., 2017; Bublatzky et al., 2014; Wieser et al., 2014), we investigated ERP differences in source recognition. Similar to previous research (Wilding et al., 1995, Wilding & Rugg, 1996; Henson et al., 1999; Jaeger et al., 2009; Mangels et al., 2001), we specifically focused on trials in which both item and source recognition were correctly classified. Here, pronounced late negativities emerged (central 420-520ms and parieto-occipital 716-900ms) for old faces that have been encountered previously within a threatening context. Despite the absence of conscious source recognition, source information from the encoding phase modulated face processing in the subsequent recognition trials. Thus, threatening context information seems to influence both early perceptual and later elaborate processing stages. These findings suggest that a threatening environment influences subsequent perception and attention allocation even without conscious representation. In other words, a specific person might appear dangerous (and be avoided) even without knowing why. A broad range of mental disorders (e.g., Social Anxiety Disorder, Posttraumatic Stress Disorder) are accompanied by such unspecific and vague feelings of threat which could be the result of previous threatening encounters which cannot be remembered.

Implications, limitations, and future directions. Several noteworthy aspects of the present study need to be emphasized. First, in the present design, source information was more important than the target item per se. Specifically, our source manipulation emphasized the contextual colors as signals for threat-of-shock or safety, thus providing important signal value for adaptive behavior (e.g., avoidance or coping strategies). In contrast, the combined item- and source-recognition task targeted the identity information of a face (i.e., item) as prerequisite for recollecting the source information which was emotionally less relevant (i.e., displaying neutral facial expression). Future research should account for the emotional (im-)balance of item and source information. For instance, strengthening the association and integration of item and sources may boost familiarity processes (Diana et al., 2011; Kahn et al., 2004) and lead to high priority processing of compound emotional item-context information (e.g., Mather & Sutherland, 2011; Ventura-Bort et al., 2016b). Second, the present study used an incidental learning approach to test the effects of implicit perceptual processes on face and person memory. To this end, no mentioning of a memory test was made before the beginning of the encoding phase (incidental learning). While participants were instructed to watch all pictures presented on the screen, we cannot exclude attentional competition and distraction between face and context information (Schupp et al., 2007; Schupp et al., 2008). However, as we observed threat-effects during the recognition phase (i.e., in the absence of contextual threat/safety signals), successful face–context binding emerged on the neuronal but not behavioral level. Future research using incidental learning may implement direct measures of the attentional focus (e.g., using eye-tracking or online face ratings during picture viewing). Third, the utilized presentation features (i.e., 1 s picture presentation with constantly changing background colors) were apparently too difficult to enable solid memory encoding of face identity. Thus, to follow-up on the impact of threat on item- and source memory, future research may improve the quality of memory and binding of both item and source information (e.g., increasing encoding times and blocked context manipulations; Wilding, 1999; Weymar et al., 2014; Bublatzky et al., 2010). Fourth, rating and ERP differences were observed for threat versus safety context during the encoding session, thus confirming the successful threat-of-shock manipulation. However, baseline and/or online ratings of the context conditions (e.g., regarding shock expectations) would have enhanced interpretability of the instructed threat effects. Whereas previous research reported rather persistent effects of threat anticipation on self-report, peripheral and ERP measures (Bublatzky et al., 2010, 2012, 2013, 2014), potential habituation and/or extinction learning processes cannot be examined within the short time period of the encoding session (~1min 40s).

Finally, from a clinical perspective, the present data provide first indication that trait measures of social anxiety modulate threat-effects during face encoding and recognition, even at a non-conscious level. Specifically, socially anxious participants showed amplified processing of threatening (relative to safe) face–context compounds, which was not observed in less anxious participants. Thus, aversive anticipations can bias perceptual processing of item/context information and in turn modulate the perception of otherwise safe persons or situations (Wieser et al., 2014). Here, the threat-of-shock paradigm may serve as a laboratory analog to examine the anticipatory aspects of anxiety – without the actual experience of aversive stimuli or situations – which are highly relevant for various anxiety and stress-related disorders (Robinson et al., 2013; Bublatzky et al., 2014; Craske et al., 2012). Awaiting replications of our correlational findings, future research may examine item/source memory (deficits) in patients suffering from social anxiety or post-traumatic stress disorder. Here, the implementation of threatening and stressful situations may be particularly helpful to model real-life anticipatory distress, without experiencing aversive events.

Conclusion. Taken together, the present study shows that contextual threat boosts electrocortical activity during the encoding and recognition of faces. Faces that were encoded during threat-of-shock were associated with selective neural processing patterns, as revealed by enhanced early parieto-occipital and late fronto-central negative potentials compared to safety faces. Regarding face recognition, the overall memory performance for recognizing face- and source-information was relatively poor. Nevertheless, brain activity differentiates old from new faces as indicated by an enhanced fronto-central positivity for previously encountered faces. Moreover, old faces presented within a context of threat were associated with enhanced late central and parieto-occipital negativities compared to old safe faces. These findings suggest selective attention to faces, which have previously been encountered in a threatening environment; intriguingly, these memory effects are not reflected in behavioral recognition performance. Thus, contextual threat amplifies the processing of social information even in the absence of conscious recognition. Future research may help to clarify clinical implications of threat-biased face/person memory in stress-related disorders.

Chapter III

Study 2: Social learning in individuals with adverse childhood experiences: Memory processes as a function of instructional and observational threat and safety learning.

An adapted version of this chapter has been submitted to the European Journal of Psychotraumatology and is currently under review as "Schellhaas, S., Schmahl, C., & Bublatzky, F. (under review). Social learning in individuals with adverse childhood experiences: Memory processes as a function of instructional and observational threat and safety learning."

3.1 Abstract

Adverse childhood experiences (ACE) are often associated with stress and anxiety-related disorders in adulthood and learning and memory deficits have been suggested as a potential link between ACE and psychopathology. In this preregistered study, the impact of social threat learning on the processing, encoding, and recognition of unknown faces as well as their contextual settings was measured by recognition performance and event-related brain potentials (ERP). Sixty-four individuals with ACE encoded neutral faces within threatening or safe context conditions. During recognition, participants had to decide whether a face was new or had been previously presented in what context (item-source memory), looking at old and new faces. For visual working memory, participants had to detect changes in low and high load conditions during contextual threat or safety. Results showed a successful induction of threat expectation in persons with ACE. In terms of face and source recognition, overall recognition of safe and new faces was better compared to threatening face-compounds, with more socially anxious individuals having an advantage in remembering threatening faces. For working memory, an effect of task load was found on performance, irrespective of threat or safety context. Regarding electrocortical activity, an old/new recognition effect and threat-selective processing of face–context information was observed during both encoding and recognition. Moreover, neural activity associated with change detection was found for faces in a threatening context, but only at high task load, suggesting reduced capacity for faces in potentially harmful situations when cognitive resources are limited. Further behavioral and ERP findings are discussed within the framework of stress-related disorders.

3.2 Introduction

Adverse Childhood Experiences (ACE) such as sexual, physical, emotional abuse or neglect in childhood or adolescence are a lifelong burden for those affected, as well as for family members and society in general. Individual consequences often include a severely reduced quality of life, which manifests itself in a wide range of health, social, financial, and occupational problems (e.g., reduced work performance, sick leave). In addition, mental health impairments, including risk for mental disorders and comorbidities, impose enormous socioeconomic costs (Dube et al., 2003). Impaired cognitive functions such as attention, perception, and memory, which focus on arousing and threatening information, are considered vulnerability factors and precursors for the development and persistence of psychopathology (Bar-Haim et al., 2007).

Understanding the deleterious effects of stress has been the subject of much neuroscience research. Severe stressful events can lead to prolonged and/or blunted response of the hypothalamus–pituitary–adrenal axis (McEwen, 2002), affecting brain morphology involved in learning and memory (e.g., hippocampus, prefrontal cortex; Bremner et al., 2003) and providing the basis for poor encoding of stressful situations. Moreover, stress-related neuronal and humoral responses have been linked to changes in information processing. For instance, biased attention and memory for negatively arousing information are associated with cognitive impairments such as distractibility or concentration problems (e.g., negativity bias;

Letkiewicz et al., 2020). Here, measurements of electrocortical activity provide high-resolution insights into the temporal dynamics of perceptual and attentional processes. Several components of event-related brain potentials (ERP) have been identified as sensitive to early stress, trauma experiences, and post-traumatic stress disorder (e.g., N2, P200, P300; Karl et al., 2006). For instance, individuals with a history of abuse show reduced N2 and P300 for emotional stimuli, which presumably reflect problems in disengagement from emotional information (Letkiewicz et al., 2020). Importantly, combined attentional and learning deficits could serve as a link between adverse childhood experiences and severity of psychopathology. For instance, decreased discrimination between arousing and non-arousing cues may contribute to (over-)generalization of threat and related avoidance behaviors (Stegmann et al., 2020). Moreover, threat-selective processing patterns for several emotional valence categories have been observed in individuals with posttraumatic stress disorder (PTSD), which are similar to the processing of negative stimuli in healthy participants (e.g., enhanced P300 and late positivities; Saar-Ashkenazy et al., 2015). The temporal dynamics of cognitive processes in individuals with ACE are therefore likely to reveal changes in the perception, recognition, and memory of threatening situations.

The impact of anxiety and arousal on memory processes. Anxious arousal has long been thought to modulate memory processes, and both facilitative and detrimental effects have been observed. In this regard, the fit of arousing conditions, memory system (e.g., short-term vs. long-term memory) and task characteristics (e.g., load) appear to critically modulate memory performance (Robinson et al., 2013; Moran, 2016). On the one hand, arousal can direct attention to salient stimuli and strengthen their mental representation (Mather & Sutherland, 2011). For instance, a memory-enhancing effect has been found for neutral and emotionally arousing items under arousing conditions, with weakened recall of the associative context in short-term memory tasks (Ventura-Bort et al., 2016b; Kensinger, 2009). On the other hand, arousal consumes memory resources and therefore reduces the ability to actively suppress distracting information, resulting in a detrimental effect of arousal when greater recruitment of cognitive resources is required to maintain task performance (Eysenck et al., 2007; Moran, 2016). For instance, social anxiety as an arousal factor was associated with regular capacity of visual working memory only in the absence of task-irrelevant distractors and only in low-demand tasks (Moriya & Sugiura, 2012). Moreover, focusing on short-term and source memory in healthy participants, we found that contextual threat facilitated perceptual processing for both the central (neutral) item and peripheral (arousing) context information, but it did not affect recognition performance for neither (Schellhaas et al., 2020). Thus, the interplay between arousing conditions arising from the task itself or inter-individual differences (trait anxiety, ACE) is poorly understood, and further studies are needed to examine different memory systems.

The role of social threat and safety learning in face perception and recognition. Anxious arousal can be triggered not only by one's own experiences but are also mediated by social cues in the environment. Specifically, humans learn vicariously by observing aversive experiences of others and through verbal communication (Haaker et al., 2017; Robinson et al., 2013). Although such social learning is considered extremely relevant in the development and maintenance of stress-related psychopathology (Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM-5; American Psychiatric Association [APA], 2013), only few studies have experimentally addressed this notion. Here, we examined social learning mechanisms to induce anxious arousal, focusing on person perception and recognition. Recognition of unfamiliar faces is generally poor (Burton & Jenkins, 2011), but arousing contextual features affect both perception and neural processing of these faces and the likelihood of recognizing them or the contextual settings. This effect is modulated by interindividual differences, for instance, individuals with depression showed enhanced P300 amplitudes for social threat information (Iffland et al., 2021). However, whether ACE in conjunction with contextual threat alters memory for faces and contextual information is poorly understood.

Study objective and hypotheses. This pre-registered study (https://osf.io/gzpev) examined the influence of socially learned threat and safety on face memory, context memory, and visual working memory in participants with ACE as a potential mediator of psychopathology. To this end, an item/source memory and a change detection task were performed with unknown faces as task stimuli (Sessa et al., 2011; Weymar et al., 2013). Importantly, both tasks were completed within contextual settings serving as signal for shock threat or safety, and threat/safety associations were learned vicariously through observing others or by means of verbal instructions. Based on previous research on observational and instructional learning (Olsson & Phelps, 2007; Bublatzky, Guerra, & Alpers, 2020), we predicted pronounced valence, arousal, and threat ratings for threat context, and threat-selective electrocortical processing as indicated by early and late parieto-occipital and fronto-central negativity relative to instructed/observed safety conditions (Schellhaas et al., 2020). No differences between instructional and observational learning were expected for any of the measures (Olsson & Phelps, 2007).

Regarding item/source memory performance, enhanced recognition was expected for faces from a threat context with a possibly impaired recognition ability for the context itself (Ventura-Bort et al., 2016b; Bisby et al., 2018). Visual working memory performance depends on the availability of cognitive resources and should be better if only one target stimulus is presented in comparison to two (i.e., different load conditions). Regarding the impact of threat, previous findings are mixed showing either reduced ability to actively inhibit the threatening information taking away attentional resources (Moran, 2016), null effects (Ward et al., 2020), or even better recruitment of cognitive resources (Moriya & Sugiura, 2013). Regarding electrocortical processing, we predicted differential processing of faces during threat relative to safe context conditions during the encoding session of the item-source memory task (N170, EPN, LPP; Schellhaas et al., 2020; Bublatzky et al., 2020). For face recognition, we expected an old/new recognition ERP effect with enhanced positivity for previously presented faces in earlier (~300-500ms parietal-occipital) and, depending on correctly assigned threat context, in a late time window (~600-800ms fronto-central; Schellhaas et al., 2020; Rugg & Curran, 2007). For visual working memory capacity, indexed by the contralateral delay activity (CDA), an increased amplitude was expected when the number of objects maintained in WM increases (Ikkai et al., 2010) and while a threat context as task irrelevant disrupting information is present (Ward et al., 2020). In addition, the N170, a component that is reliably elicited by facial stimuli (Schindler & Bublatzky, 2020), was exploratively analyzed as an index of cognitive load and face perception in the working memory task (Morgan et al., 2008).

Regarding the impact of ACE, we hypothesized that higher levels of traumatization were associated with increased neural processing of the threat but also the safe condition (especially at moderate-severe levels; Karl et al., 2006). It was therefore predicted that differential processing of threat and safety conditions would decrease, as would the overall strength of processing (Galletly et al., 2001). Regarding behavioral measures, we assumed that perceived threat affects cognitive processing and memory and tends to inhibit adequate performance (Pechtel & Pizzagalli, 2011). Therefore, retrieval of face–context compounds under threat, as well as retention of these associations in working memory (Goodman et al., 2019; especially under high load), was expected to be worse with increasing ACE levels. Although a dimensional approach was pursued, group comparisons were made between low, medium, and high levels of anxiety and ACE, and their subgroups, because some effects of early trauma on cognition may depend on type and severity (Herzog & Schmahl, 2018).

3.3 Methods

Participants

Sixty-four participants (59 females) between the age of 19 to 60 years (M = 32.45, SD = 11.05) were recruited from all over Germany. As an inclusion criterion, all participants had experienced adverse childhood experiences (from the age 0 to 18, abuse and neglect) as determined by the German version of the Childhood Trauma Questionnaire (CTQ; German Version, Bernstein et al., 1998). In a pre-screening, one item of every subscale (sexual abuse, physical neglect and abuse, emotional neglect and abuse) was presented with an overall value of ≥ 1 as cut-off criteria for participation. Sample size was determined by power analysis based on previous findings, detecting medium effects (Sessa et al., 2011) and statistical power of $1-\beta \geq .80$ (g*power, Erdfelder et al., 1996). Exclusion criteria were acute and/or chronic physical diseases (e.g., cardiovascular, respiratory, or neurological diseases), psychotic disorders, use of psychotropic drugs (except selective serotonin and norepinephrine reuptake inhibitors [SSRIs and SNRIs]), and current (past 12 months) substance dependence and/or abuse. Inclusion/exclusion criteria were verified by means of an interview prior to participation which also served as diagnostics acquisition (Structured Clinical Interview for DSM-IV (APA, 2000), SCID-I (36 participants) or DSM-5 (APA, 2013), SCID-5 (28 participants)), done by PhD students trained in conducting psychological diagnostics (see supplements). For the diagnostics, data of five participants is missing due to technical errors in data collection and storing.

Participants were recruited via advertisements placed on various homepages (Central Institute of Mental Health Mannheim [CIMH]; Research Training Group GRK 2350), in local newspapers, and flyers at the CIMH and psychotherapist offices. Ethics approval was given by the local ethics committee and participants provided written informed consent to the study protocol and received monetary compensation (30€).

Materials and memory tasks

Participants performed two experimental tasks, a combined item/source memory (ISM) and visual working memory (VWM) task, both using face stimuli. Neutral face pictures of in total 150 actors (half females; 90 were chosen for the ISM task and 60 for the VWM; see supplements) were selected from the Karolinska Directed Emotional Faces (Lundqvist et al., 1998), the Radboud (Langner et al., 2010), and the NimStim database (Tottenham et al., 2009). Pictures (442×606 pixels) were transformed into grayscale, normalized for brightness, cropped with an elliptic mask to remove hair and ears, and a black frame was used to replace the original background (Adobe Photoshop CS2). For both tasks, face pictures were presented with contextual colors as a backdrop (1280×1024 pixels), using different color combinations, either blue and green (RGB values: 0,255,0 and 0,0,255), or red and yellow (RGB values: 255,0,0 and 255,255,0); color assignment to condition was counterbalanced.

Item/source memory task. Participants were asked to recognize a previously presented face and to recall the specific contextual features of the face (Schellhaas et al., 2020). To this end, a random selection of 60 out of 90 face pictures was chosen, which were surrounded with two different colored background frames (e.g., blue and green), 30 pictures for each color. Participants' task was to look carefully at the faces and memorize them while being instructed to recognize them later (explicit learning instruction for item memory). The additional source identification task was not mentioned (implicit learning of context information).

During an encoding session, the 60 pictures were presented for 6 seconds each (‘old’ faces), separated by an inter-trial interval (ITI) ranging between 620 and 1150ms, and the frame colors alternated in blocks of 10 pictures (i.e., 3 green or red and 3 blue or yellow blocks). Without delay, the recognition session started, and all 60 old faces (without color frames) were presented intermixed with 30 additional new pictures in randomized order. Participants’ task was to indicate the contextual color against which a face had been presented in the encoding phase (i.e., combined item and source memory), or whether it was a new face. Pictures remained on the screen until the participant responded by pressing one of three keyboard buttons. The behavioral options were either a blue (yellow) or a green (red) button (corresponding to the background, classified as old threat and safety faces), or a white button for newly presented pictures (new faces). If participants recognized a picture as being from the encoding phase but did not remember the context, they were instructed to guess the context. There was no time limit for responding and no feedback on accuracy was provided. Each choice was followed by a 1s ITI showing a black rectangle replacing the pictures and the next face was presented (Figure 6).

Visual working memory task. A change detection task served to examine visual working memory (Stout et al., 2013; Sessa et al., 2011). Each trial consisted of a memory and a test array displaying 2 or 4 faces simultaneously (Figure 6). Before the memory array was presented, two arrows (pointing 200ms to the left or right) indicated the position of the target faces, which were the face(s) to be remembered. The memory array was presented for 500ms and was followed by an empty retention interval (i.e., no faces) screen for 900ms in the same color as the memory array. Following, the test array was presented and the participants had to decide whether (one of) the target face(s) had changed the identity by pressing one of two buttons (stating “the same” or “different”). Regardless of the overall number of faces in the memory array (2 vs. 4) only one face`s identity was changed. The change detection task was performed with two different background colors (either blue and green or red and yellow), and this context condition alternated every 10 trials.

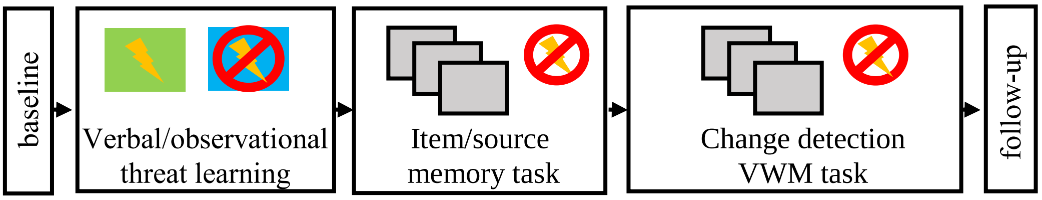
Participants’ task was to focus on the arrow indicated side and memorize only the face(s) in the memory array presented on this side. The number of target faces served to manipulate the task load (low vs. high load), and varied randomly within each context condition. In half of the trials, the faces on the memory and test array were identical, for the other trials one face on the arrow-cued side of the memory array was replaced with a different same-gender face in the test array. There was no response time limit and after the response, a fixation cross was presented (ITI 500ms) indicating the start of the next trial. Participants completed 16 practice trials without contextual colors (two for each combination of high- vs. low-load and changed vs. non-changed target). In the test phase, participants completed 6 runs without breaks in each color, each with 64 trials (768 trials in total), presented in an evenly changing order and evenly distributed for high/low load and change/no change trials (Figure 6). For both tasks, the change of the background color was indicated by the presentation of a picture of the new color without a face for 4s.

Stimuli were presented on a 22-inch computer screen placed approximately 1m in front of the participants using OpenSesame software (Mathôt et al., 2012).

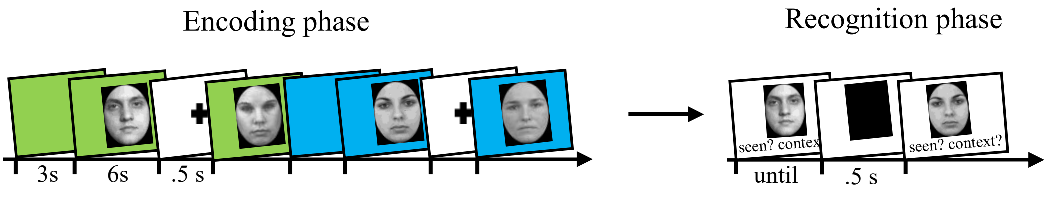
Procedure

A first set of questionnaires was completed approximately 1-2 days before the testing day (CTQ and SCID). Upon arrival in the lab, the EEG sensor net was attached and questionnaires on social anxiety, depression and state anxiety were completed (Social Phobia Inventory, SPIN, Stangier & Steffens, 2001; State-Anxiety Inventory, STAI-S, Laux et al., 1981; Table 3). Following, a baseline rating regarding the experimental background colors (blue/green, or yellow/red, serving later as threat/safety signals) was performed using valence, arousal (Self-Assessment Manikin, SAM; Bradley & Lang, 1994, on a scale from 1-9) and perceived threat scales (Likert scale from 0 to 10). To trigger aversive expectations about imminent shocks, a fake stimulation electrode was attached to the inner forearm of the non-dominant arm. Participants were then told that the expected shock intensity would be “maximally unpleasant but not yet painful” and that they would receive a maximum of three electric shocks throughout the entire experiment.

(A) Procedure



(B) Item / source memory task



(C) Visual working memory task: change detection

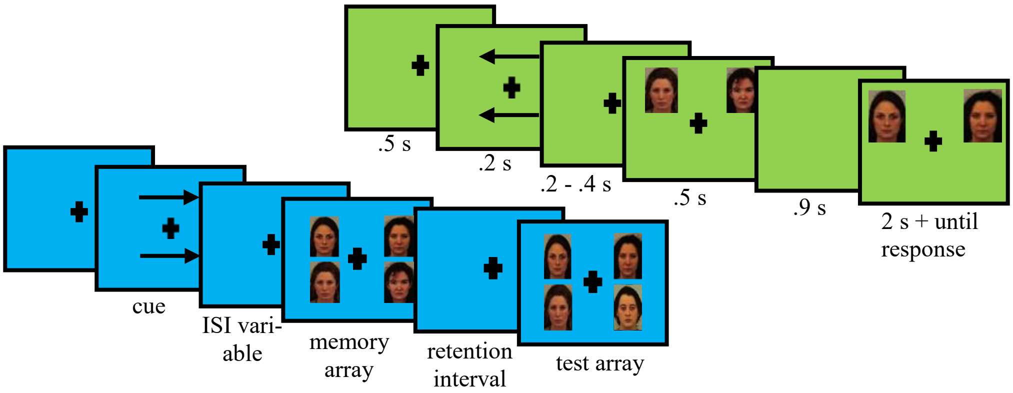


Figure 6. (A) Schematic illustration of the experimental procedure. Background colors served as verbally instructed/observed threat-of-shock or safety cues in both following memory tasks, which were performed in a randomized order. (B) In the item/source memory task, during the encoding phase, 60 pictures of male and female faces displaying neutral expressions were presented for 6 s each (variable ITI) in front of two different colors that alternated in blocks of 10 faces, each block indicated by a colored frame for 3s. Participants' task was to memorize the faces, without a mentioning of the colors. During the recognition phase, the 60 old and 30 new faces were presented intermixed without background frames. Participants' task was to decide whether the face was shown before and with which color (combined old/new item and source recognition task), or whether it was not previously presented. (C) For visual working memory task (change detection), in a memory array either four or two faces were presented for 200ms. After a 900ms retention interval, participants had to decide whether one face on the arrow indicated side (arrow cue presented for 200ms before the memory array) had changed identity or not in the test array.

Participants acquired threat and safety associations socially via either verbal instructions or observations. In the instruction group, participants were verbally instructed that specific background colors will serve as signals for shock threat (e.g., green) or safety from shocks (e.g., blue; Bublatzky, Gerdes, & Alpers, 2014; Schellhaas et al., 2020). The observation group learned threat/safety associations with contextual colors vicariously by viewing a video displaying another (fake) participant as a demonstrator (learning model) undergoing a differential threat conditioning experiment (Haaker et al., 2017).

These videos consisted of pre-recorded experimental sessions with a female and male demonstrator within the same experimental environment. Videos contained 24 trials each (edited for brightness, trial length, and 4-minutes duration using Adobe Premiere Pro X).

Two thirds of the video-threat trials were paired with a shock, indicated by the demonstrator twitching their arm and showing a frowning facial expression. The participants themselves were told that they would not receive shocks during the video; however, they should pay close attention to the demonstrator and the colors as they would receive shocks with the same color in the next part of the experiment. Participants viewed videos with a same-sex demonstrator; assignment of colors to threat/safety conditions were balanced across participants.

The two memory tasks were completed in randomized order. At the end of the experiment, the videos were rated (using Likert scale forms 0-9) and revealed a medium-high demonstrator-observer agreement regarding naturalness (M = 4.98, SD = 2.91), identification (M = 5.06, SD = 2.95), expressiveness (M = 6.94, SD = 2.34), discomfort (M = 5.78, SD = 2.71), and empathy (M = 5.60, SD = 3.06). As a manipulation check, context colors were rated regarding valence, arousal, and perceived threat following both memory tasks. Indicating the successful learning of threat/safety contingencies, all participants were able to indicate the correct threat and safety color at the end of the experiment. Finally, participants were debriefed.

Table 3. *Subjective ratings for the overall sample and split by the between factor social learning (observation/instruction).*

	SPIN	STAI-S	CTQ_{TOT}	CTQ_{sexab}	CTQ_{emoab}	CTQ_{phyab}	CTQ_{emoneg}	CTQ_{phyneg}
Tot	23.19 (13.48)	37.94 (5.32)	73.00 (23.85)	10.78 (6.76)	16.16 (6.10)	9.81 (4.87)	16.35 (6.16)	9.68 (4.29)
Obs	23.97 (15.28)	37.61 (4.36)	76.13 (20.69)	11.23 (6.32)	16.80 (5.92)	10.27 (4.36)	17.53 (6.16)	10.10 (4.14)
Inst	22.42 (11.67)	38.24 (6.13)	70.15 (26.40)	10.36 (7.20)	15.58 (6.30)	9.39 (5.32)	15.27 (6.05)	9.30 (4.46)

Note. Values represent mean (M) and standard deviation (SD); Tot = total sample; Obs = observation group; Inst = verbal instruction group; SPIN = social phobia inventory; STAI-S = State/Trait anxiety inventory – state; CTQTot = total score childhood trauma questionnaire; CTQsexab = subscale sexual abuse of CTQ; CTQemoab = subscale emotional abuse of CTQ; CTQphyab = subscale physical abuse of CTQ; CTQemoneg = subscale emotional neglect of CTQ; CTQphyneg = subscale physical neglect of CTQ.

Data recording and reduction

Electrocortical activity was recorded using a 65-channel system (BrainProducts, Munich, Germany). Ag/AgCl active electrodes were placed in a cap using a 10-10 electrode placement standard with FCz as reference electrode (Falk Minow Services, Herrsching, Germany). EEG was recorded continuously with a sampling rate of 500Hz and filtered online from 0.1 to 100Hz using Vision Recorder acquisition software and BrainAmp DC amplifiers (BrainProducts). Electrode impedance was kept below 20kΩ (manufacturer recommendation). Offline data analyses were done using VisionAnalyzer 2.0 (BrainProducts) and EMEGS (Version 2.7, Peyk et al., 2011) including conversion to an average reference, 30Hz low-pass filtering, artifact detection, sensor interpolation, and baseline correction (200ms). Artifacts were rejected from trials exceeding ±70µV (e.g., eye blinks), on in average 1.29% of the trials for the ISM and 2.37% of the trials for the VWM.

Stimulus-synchronized epochs were extracted and lasted for the ISM task from 200ms before to 1000ms after stimulus onset (encoding and recognition session), and for the VWM task from 200ms before to 1200ms after stimulus onset of the memory array (including retention and test array). For the VWM task, difference waves for the contralateral-delay activity (CDA) were computed by subtracting the average activity recorded by electrodes ipsilateral to the arrow-cued visual field of the memory array from the average activity recorded at symmetrical electrodes contralateral to the arrow-cued visual field of the memory array (i.e., CP5/CP6). The CDA is classified as activity during the retention interval (500–900ms after memory array) and thought as reflecting working memory capacity (Sessa et al., 2011). Finally, separate average waveforms were calculated for each condition, for each sensor (difference of sensors for CDA) and participant.

One participant was excluded from all EEG analyses due to technical errors during recording (no data was collected). For the EEG analyses of the VWM task, additional four participants were discarded due to poor data quality (more than 30% of trials with artifacts). Due to missing values, two participants were excluded from the arousal ratings and seven from the perceived threat ratings. One and two participants for STAI-S and CTQ respectively with lack of data were excluded from covariation analyses.

Data analysis

The Greenhouse Geisser procedure was used to correct violations of sphericity and Bonferroni correction for multiple comparisons. As a measure of effect size the partial eta squared (η_p^2) is reported. Statistical analyses were conducted using SPSS (version 25), for the MPT analyses the TreeBUGS package (Heck et al., 2018) in R studio (version 3.6.2, R Core Team, 2016). Significant effects were followed up by a separate two-tailed t-test with significance level set to $p < 0.05$. Due to our focus on adverse childhood experiences, social and general anxiety, we also included questionnaire scores of the SPIN, STAI-S and CTQ_{TOT} as well as the five subscales of the CTQ (emotional and physical abuse and neglect, sexual abuse) as exploratory covariates.

Self-report data. Separate 2×2×2 repeated measure ANOVAS were calculated for valence, arousal, and threat ratings. Within-subject factors were Time (baseline vs. follow-up) and Context (threat vs. safety), as well as the type of Learning (observation vs. instruction) as a between-subject factor.

Behavioral data. To quantify memory performance in the item/source memory task, hit rates (HR, correct responses divided by all responses in target trials) and false alarm rates (FAR, incorrect responses divided by all responses in non-target trials) as well as item recognition (HR-FAR) were computed. Source-memory for the encoding context of faces was analyzed using the average conditional source identification measure (ACSIM), which is the number of all correct source identifications divided by the overall number of faces that were identified as "old" (Bell & Buchner, 2011). For the visual working memory task, HR and FAR were also calculated and visual working memory was quantified with a standard index of sensitivity by subtracting the inverse-normal transform of the FAR from the inverse-normal transform of the HR, deriving from signal-detection theory and reflecting change detection ability ($d' = Z_{HR} - Z_{FA}$). For HR and FAR of zero a constant of .5 was added to the number of hits and false alarms and the number of detection and signal trials was increased by 1, as proposed by Hautus (1995). Additionally, we calculated the number of encoded faces in each condition using Cowan's *k* ($\underline{k}$ = load*(HR-FA)), serving as an index of VWM capacity.

All item/source memory measures were analyzed by 2×2 ANOVAs, with Context (threat vs. safety) as a within-subject and Learning (observation vs. instruction) as a between-subject factor. Visual working memory was analyzed by a 2×2×2 ANOVA, with the additional within-subject factor of memory Load (low vs. high).

Hierarchical multinomial processing tree (MPT) modelling was used to disentangle item recognition, source recognition, and guessing biases for the item/source memory task. We used the two-high-threshold model of source monitoring (2HTSM; Bayen, Murnane, & Erdfelder, 1996) with three decision trees (i.e., faces from threat or safety sources, or new faces; Figure 7), and individual questionnaire scores included in the MPT model (Arnold, Bayen, & Böhm, 2015). As a between-factor, Learning (observation vs. instruction) was included, separate parameter estimates were calculated for each of the three decision trees. For details on MPT analyses see supplementary material.

ERP data. Regarding the encoding session of the item/source memory task, statistical analyses were computed for an early time window (252–400ms) over parieto-occipital sites (PO7/PO8) and for a late time (540–800ms) window over fronto-central sensors (C1/C2/FC1/FC2). For the recognition session, we used a three-step procedure with increasing informative value (Schellhaas et al., 2020). First, for the classical Old/New recognition effect, only trials with correct item recognition were included (regardless of source identification) and old versus new faces were compared in a late time window (660–800ms) over parieto-occipital sensors (PO9/PO10). Second, only trials with correct item and correct source identification were included for advanced item-source Old/New analyses. Those were analyzed in a late time window (752–900ms) over central sensor sites (Cz/CPz/C1/C2/C3/C4). Third, trials based on correct item and source identification were included separately for old-threat, old-safe, and new faces. Here, a late time window (700–900ms) for parieto-occipital sensors (PO3/PO4/POz) was used.

For the visual working memory task, difference waves were computed between 500–900ms after the onset of the memory array at CP5-CP6 sensors. For visualization, waveforms were low-pass filtered (10 Hz). Additionally, the N170 component was computed over parieto-occipital sites (PO9/PO10), within a ±40ms window centered on the maximum peak of the grand-average means (147ms).

Separate repeated measure ANOVAs were computed that considered the experimental Context (threat vs. safety; for ISM encoding and recognition, as well as VWM difference waves), Recognition (old vs. new; for ISM recognition), Load (low vs. high; for VWM task), Learning group (observation vs. instruction; as between-subject factor), Laterality (left vs. right hemisphere).

3.4 Results

The data that support the findings of this study are openly available in OSF at https://osf.io/u2nq7/?view_only=59fd90797a184c7d8cac8ad194d20f0e

Self-report data

Both observational and instructional learning groups successfully acquired threat and safety associations. Significant interactions emerged for Context × Time (valence: $F(1,62) = 17.29$, $p < .001$, $\eta_p^2 = .23$ [.08,.35], arousal $F(1,60) = 20.50$, $p < .001$, $\eta_p^2 = .26$ [.11, .39], threat $F(1,55) = 14.14$, $p < .001$, $\eta_p^2 = .21$ [.06, .33]. After learning, the threat context was perceived as more unpleasant, arousing, and threatening compared to safety (valence*:* $F(1,63) = 25.28$, $p < .001$, $\eta_p^2 = .29$ [.15, .45], arousal*:* $F(1,63) = 37.07$, $p < .001$, $\eta_p^2 = .37$ [.21, .49], threat: $F(1,60) = 32.60$, $p < .001$, $\eta_p^2 = .35$ [.19, .48]). Interestingly, no differences were found between learning groups (*F*s < 1.48, *p*s > .23), and no interactions Learning × Context emerged (*F*s < 1.14, *p*s >.29). Moreover, ratings did not co-vary with SPIN, STAI-S and CTQ_{TOT}. A significant positive correlation emerged between social anxiety and trauma scores (SPIN × $CTQ_{TOT,}$ $r(60) = .28$, $p < .05$).

Recognition performance

Item and source memory (ISM). Conventional measures of recognition performance revealed no effects of aversive anticipation on item (HR-FAR; Context $F(2, 62) < 1$) or source identification [ACSIM; Context $F(1,64) = 1.40$, $p = .24$; $HR_T = .71$ ($SD = .16$), $FAR_T = .23$ (SD = .15), $HR_S = .72$ ($SD = .14$), and $FAR_S = .23$ ($SD = .15$). No differences emerged between observational and instructional learning (*F*s < 1.51, *p*s > .22). Regarding questionnaire data, a covariation of the HR with the CTQ subscale sexual abuse and emotional neglect was found, $F(1,62) = 4.83$, $p < .05$, $\eta_p^2 = 0.08$ [.00, .19] and $F(1,62) = 4.31$, $p < .05$, $\eta_p^2 = .07$ [.00, .18], indicating that the HR of highly traumatized individuals increased during threat.

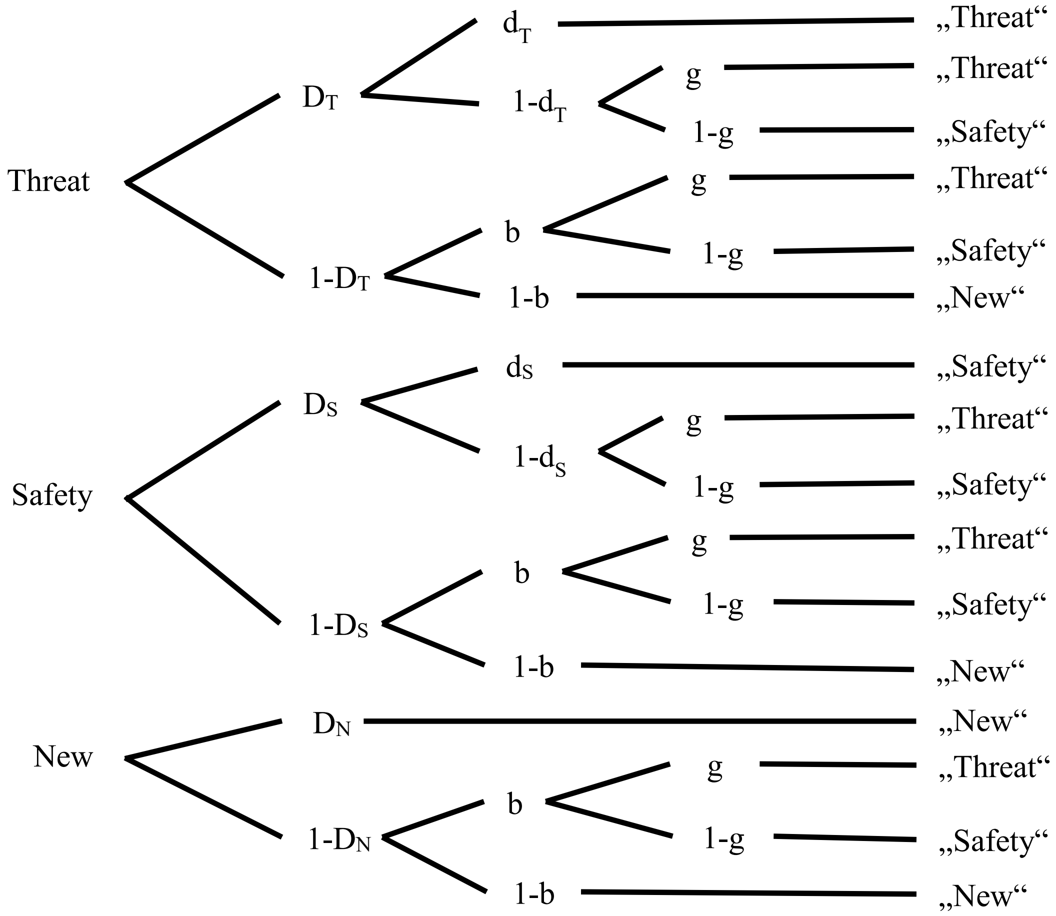


Figure 7. Submodel 5d of the two-high-threshold model of source monitoring. The model parameters represent transition probabilities between latent cognitive states; D_T = probability of detecting that an item is old from a threatening source; D_S = probability of detecting that an item is old from a safe source; D_N = probability of detecting that an item is new; d_T = probability of correctly remembering the shock source of an item; d_S = probability of correctly remembering the safe source of an item; g = probability of guessing that an item is from the shock source; b = probability of guessing that an item is old. Adapted from Arnold et al., 2013; originally by Bayen et al., 1996.

Only trend-level covariations emerged between state anxiety and hit rates (STAI-S × Context, $F(1,62) = 2.68$, $p = .07$, $\eta_p^2 = .05$ [.00, .14]), as well as social anxiety and false alarm rates (SPIN × Context: $F(1,62) = 3.47$, $p = .07$, $\eta_p^2 = .05$ [.00, .16]). Specifically, in individuals with more state anxiety higher hit rates were found for safety items, and a lower false alarm rates was found for threat items in more socially anxious participant, suggesting a better performance for highly anxious persons in a safe context.

The hierarchical MPT model provides separate estimates for item and source recognition as well as the guessing parameters as a function of contextual settings (Table 4). While overlapping Bayesian confidence intervals (BCIs) indicate a non-meaningful difference, non-overlapping BCIs were observed between the recognition parameters for faces from the threat context (D_T) and safety context/new faces (D_S/D_N). In other words, faces from a threat context were recognized less well than faces from a safe source or new faces. Regarding source recognition, overlapping and very large BCIs for the threat and safety sources (d_T and d_S) indicate no difference. Moreover, given a 2:1 ratio of old to new items, there was a slightly conservative tendency towards guessing that a face was new (b = .58). The source guessing parameter (g = .51) reflects the event probability that a face was from a threatening or safe context (1:1 ratio), thus at chance level. Finally, no differences were observed between the learning groups, and no covariation emerged with questionnaires. Model fit was assessed using Klauer´s (2010) test statistic T1. The corresponding p-value was .39, indicating good model fit.

Visual working memory (VWM). A summary of the VWM performance measures is reported in Table 5. No effects of contextual threat/safety or learning type were observed for hit rate, Fs(1,60) < 1, ps > .40. However, covariations between Context and the CTQ subscales emotional abuse and emotional neglect emerged for the hit rate, $F(1,54) = 6.74$, $p < .05$, $\eta_p^2 = .11$ [.01, .22] and $F(1,54) = 5.68$, $p < .05$, $\eta_p^2 = .10$ [.01, .23]. These effects indicate that individuals with higher levels of emotional abuse had a higher chance of detecting change under threat, whereas highly emotionally neglected individuals performed better under safety than threat. High load reduced hit rate indicating better performance for a set size of 2 relative to 4 faces, Load *F*(1,60) = 166.18, *p* < .001, η_p^2 = .74 [.63, .79], the interaction of Context × Load, *F*(1,60) < 1, *p* = .58 was not significant.

Table 4. *Mean parameter estimates of the latent-trait MPT model for the recognition performance of the item/source memory task.*

Parameter	M [95 % BCI]	SD
g	0.51 [0.49 – 0.53]	0.01
b	0.58 [0.52 – 0.64]	0.03
d_T	0.23 [0.00 – 0.83]	0.23
D_T	0.16 [0.06 – 0.26]	0.05
d_S	0.04 [0.00 – 0.17]	0.05
$D_S = D_N$	0.34 [0.29 – 0.38]	0.02

Note. For the group-level estimates, posterior means (and SDs) are shown. BCI = Bayesian confidence interval. D_T, D_S and D_N =item recognition parameters, d_T, d_S = source memory parameters and b, g = guessing probabilities.

For the false alarm rate (FAR), aversive anticipation and memory load did not play a role (*F*s < 1, *p*s > .75). However, FAR was overall higher in the instructional learning group compared to the observational learning group, Learning *F*(1,60) = 5.13, *p* < 0.05, η_p^2 = 0.08 [.01, .20]. A covariation of Load and emotional and physical neglect was found, *F(*1,54) = 5.50, *p* < .05, η_p^2 = .09 [.01, .21] and *F*(1,54) = 4.92, *p* < .05, η_p^2 = .08 [.00, .21], showing a higher FAR during high load for neglected participants. A trend level effect was found Context × SPIN (*F*(1,59) = 3.71, *p* = .06, η_p^2 = .06 [.00, .19]), showing that the more socially anxious the participants were, the lower the FAR was in recognizing changes in threatening contexts.

For d‘ values (reflecting change detection ability), no main effects of threat/safety context or learning occurred, Context and Learning: *F*s(1,60) < 1.35, *p*s > .25. However, high load reduces the recognition of changes, Load $F(1,60) = 99.66$, $p < .001$, $\eta_p^2 = .63$ [.49, .71], irrespective of context conditions, Context × Load $F(1,60) < 1$, $p = .75$. Also for the d‘ parameter, a significant interaction of context and social anxiety (SPIN) was found, $F(1,59) = 4.96$, $p < .05$, $\eta_p^2 = .08$ [.00, .20]. Socially anxious participants were better in detecting a change within threatening compared to safe contexts.

Table 5. *Behavioral performance measures for the visual working memory task (change detection), derived from signal detection theory.*

	low load		high load	
	Threat	Safe	Threat	Safe
HR	0.79_1 (0.20)	0.79_1 (0.19)	0.52_1 (0.11)	0.51_1 (.11)
FAR	0.23_1 (0.15)	0.24_1 (0.15)	0.24_1 (0.17)	0.24_1 (0.15)
HR – FAR	0.55_1 (0.27)	0.55_1 (0.28)	0.28_1 (0.18)	0.27_1 (0.14)
d'	1.80 (1.03)	1.76 (1.01)	0.85 (0.61)	0.84 (0.50)
K	1.10 (0.55)	1.10 (0.55)	1.12 (0.73)	1.09 (0.56)
RT	931.32_2 (793.99)	872.05_2 (531.31)	1123.84_2 (778.96)	1132.77_2 (828.57)

Note. Values represent mean (M) and standard deviation (SD); HR = hit rate; FAR = false alarm rate; HR – FAR = recognition rate; d‘ = change detection ability; k = capacity index; RT = reaction time; $_1$Numbers represent percentages (and SD); $_2$in ms.

As indicated by k, VWM capacity was neither modified by Context, Learning, Load, nor Context × Load, *F*s < 1, *p*s > .40. Interestingly, a significant interaction of Context × Learning emerged, $F(1,60) = 5.21$, $p < .05$, $\eta_p^2 = .08$ [.01, .20], indicating that k was higher for the threatening compared to the safe context but only in the observational learning group. Moreover, there was a significant interaction of Load × SPIN, $F(1,59) = 4.18$, $p < .05$, $\eta_p^2 = .07$ [.00, .18] as well as Load × Emotional Neglect, $F(1,54) = 4.90$, $p < .05$, $\eta_p^2 = .08$ [.00, .20], and Load × Physical Neglect, $F(1,54) = 4.28$, $p < .05$, $\eta_p^2 = .07$ [.00, .20]. These interactions indicate that participants with high levels of social anxiety and emotional neglect had a higher k for the low compared to the high load condition, whereas physically neglected individuals showed a higher k for high load. Additionally, a trend level interaction emerged for Context × SPIN, $F(1,59) = 3.35$, $p = .07$, $\eta_p^2 = .05$ [.00, .17] and Context × Emotional Abuse, $F(1,54) = 5.10$, $p < .05$, $\eta_p^2 = .09$ [.01, .20], pointing towards highly anxious participants having lower VWM capacity for threat compared to safety conditions, whereas emotionally abused individuals seem to benefit more from the threat with a higher k.

Event-related potentials

Item and source memory: Encoding session. Confirming successful threat induction, differential processing of threat compared to safety contexts was observed during the encoding session (Schellhaas et al., 2020; Figure 8), Context $F(1,61) = 8.33$, $p < .01$, $\eta_p^2 = .12$ [.02, .27], with more pronounced parieto-occipital positivities (252-400ms) over the right hemisphere, $F(1,61) = 7.56$, $p < .01$, $\eta_p^2 = .11$ [.02, .24]. As in our previous study, a trend-level interaction Context × SPIN show increased differences between threat and safety in more socially anxious participants, $F(1,61) = 3.37$, $p = .07$, $\eta_p^2 = .05$ [.00, .16]. Moreover, a later threat-enhanced negativity emerged over fronto-central sites (540-800ms), Context $F(1,62) = 3.13$, $p = .08$, $\eta_p^2 = .05$ [.00, .16], with more pronounced amplitudes over the right hemisphere, $F(1,62) = 6.07$, $p < .05$, $\eta_p^2 = .09$ [.01, .21]. The observation and instruction group did not differ, Learning $F < 1$.

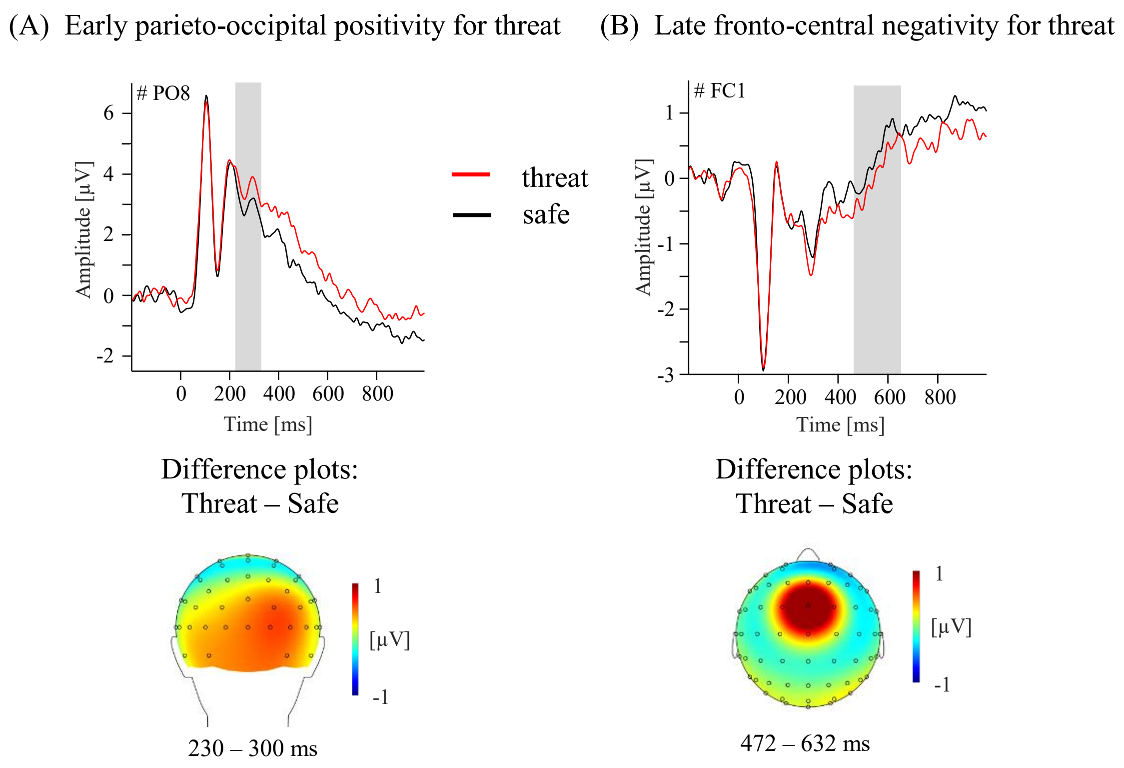


***Figure 8*. Item/source memory encoding phase: Threat effect**. Illustration of the main effect Context in the encoding session. (A) Grand averaged ERPs prompted by faces presented with threat-of-shock or safety for an exemplary parieto-occipital sensor and topographical difference maps (threat – safe) displaying the averaged time interval (230–300 ms) plotted on the back of a model head. (B) Grand averaged ERPs for faces presented within threat-of-shock or safety for an exemplary fronto-central sensor and a topographical difference map (threat – safe) displaying the averaged time interval (472–632 ms) plotted on a top view of a model head.

Item and source memory: Old/new and context recognition. In a first step, we quantified the old/new effect using trials with correct face recognition only (i.e., regardless of correct or incorrect source allocation). Correctly identified old trials were associated with a more pronounced negativity over parieto-occipital sites (600-800ms), $F(2,56) = 4.50$, $p < .05$, $\eta_p^2 = .07$ [.00, .18]. This effect directly replicates our previous study in healthy participants (Schellhaas et al., 2020). No effects were observed for Learning or Laterality, Fs(2,56) < 2.75, ps > .10.

In a second step, trials were selected based on correct item recognition. As in previous research, old compared to new faces were associated with an enhanced positivity over central sensor sites (752-900ms), $F(2,56) = 3.60$, $p = .06$, $\eta_p^2 = .06$ [.00, .18], (Figure 9), reflecting in parts our behavioral findings. No effects of Learning or Laterality occurred, *F*s < 2.30, *p*s > .11.

Third, differential processing based on correct context recognition (relative to identifying a face as new) was associated with late parieto-occipital positivity (700-900ms), Context $F(2,57) = 3.19$, $p < .05$, $\eta_p^2 = .05$ [.00, .22]. Post-hoc tests showed enhanced positivity for faces originating from a threat compared to a safe context or new faces (Figure 9). Enhanced amplitudes were found on the right hemisphere, Laterality $F(2,57) = 6.30$, $p < .01$, $\eta_p^2 = .10$ [.04, .31]. Learning groups did not differ, $F < 1$.

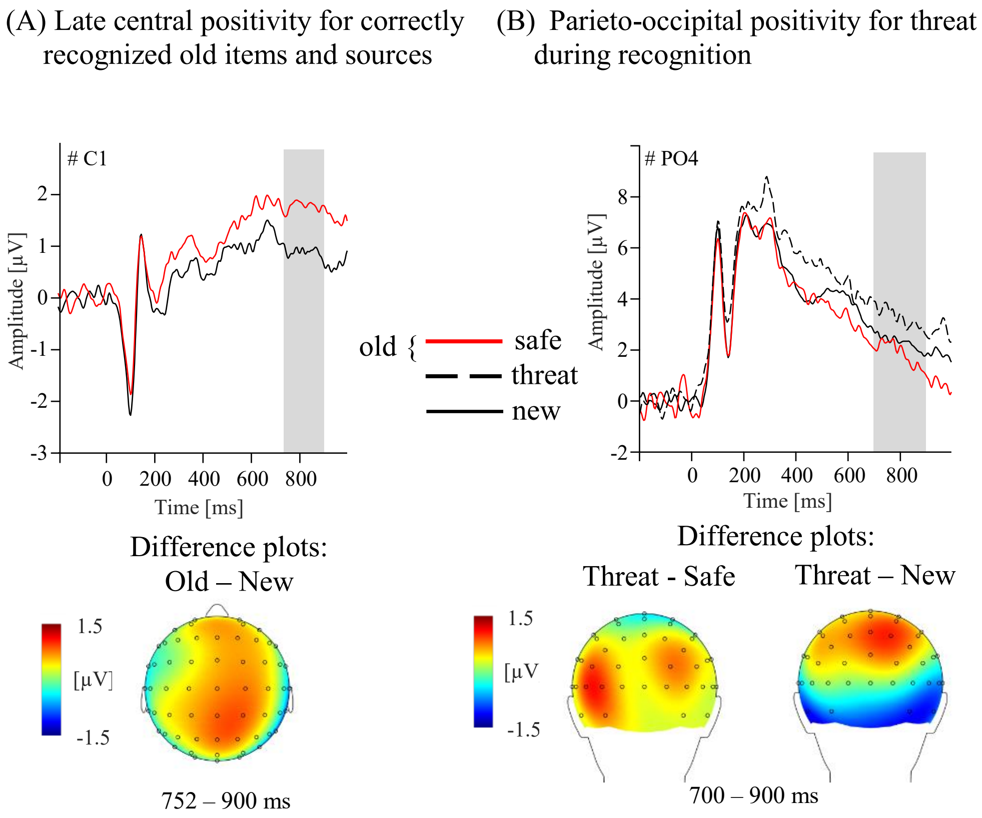
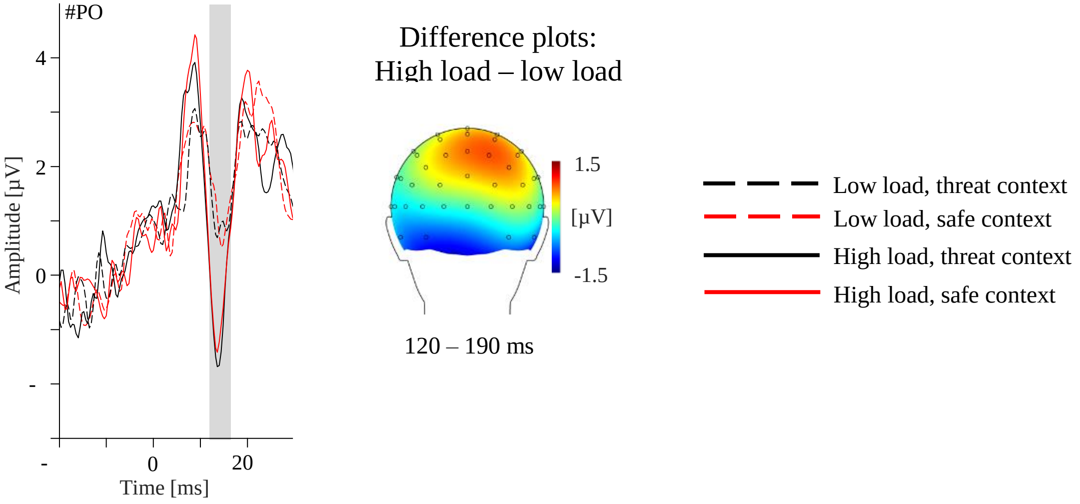


Figure 9. **Item/source memory recognition phase: Old/New and Context effect**. Illustration of the old/new effect in the recognition session (A). Grand averaged ERPs prompted by correctly classified old faces with correct source allocation, and new faces for an exemplary central sensor and a topographical difference map (old – new) displaying the averaged time interval (752 – 900 ms) plotted on a top view of a model head. Illustration of the main effect of Context in the recognition phase (B). Grand averaged ERPs for faces which have been previously presented within threat-of-shock or safety in the encoding session and new faces for an exemplary parieto-occipital sensor and topographical difference maps (threat – new, threat -safe) displaying the averaged time interval (700–900 ms) plotted on a back view of a model head.

Visual Working Memory: N170 and contralateral delay activity. A more pronounced N170 was found for high- compared to low-load over parieto-occipital sites, Load $F(1,57) = 21.62$, $p < .001$, $\eta_p^2 = .28$ [.25, .54]. There was no main effect of Laterality, Learning or Context, Fs < 1.49, ps > .23 (Figure 10).

A main effect of Context emerged for CP1-CP2 (236–336ms), $F(1,57) = 6.61$, $p < .05$, $\eta_p^2 = .10$ [.01, .24], showing an pronounced negative CDA for the safe compared to threat context, indicating that the safe context gained more access into working memory than the threatening context (Figure 10). Unexpectedly, there was no main effect of Load, *F*(1,57) = 1.19, *p* = .28, η_p^2 =.02 [.00, .11], or interaction Context × Load, *F*(1,57) = 2.70, *p* = .12, η_p^2 =.05 [.00, .16], suggesting that working memory capacity was not modulated by task difficulty and memory strain, even with changing context conditions. No main effect of Learning type occurred, *F* < 1. However, follow up tests for the significant Context × Learning interaction, *F*(1,57) = 9.63, *p* < .01, η_p^2 = .14 [.03, .28], revealed that the observation group had particular difficulty retaining the safety information from decreasing working memory capacity indicated by an enhanced negative CDA for the safety compared to the threat context. There was no difference in the threat compared to the safe context for the instruction group. Laterality had no impact, *F*s < 1. A marginal context effect emerged for CP6-CP5 (500-900ms), *F*(1,57) = 3.72, *p* = .059, η_p^2 = .06 [.00, .18], with an enhanced CDA amplitude for the safe compared to the threat context. In this later time window, an enhanced CDA was observed for high- relative to low-load condition, *F*(1,57) = 5.81, *p* < .05, η_p^2 = .09 [.01, .22], and this finding is in accordance with the hypotheses that higher load consumes working memory resources and reduces its capacity. No interaction effect between Context × Load, nor a main or interaction effect of learning type occurred, *F*s < 3.72, *p* > .059.

(A) Enhanced N170 (parieto-occipital) for the high load condition regardless of context



(B) Enhanced CDA (central-parietal) for threat only in the high load condition

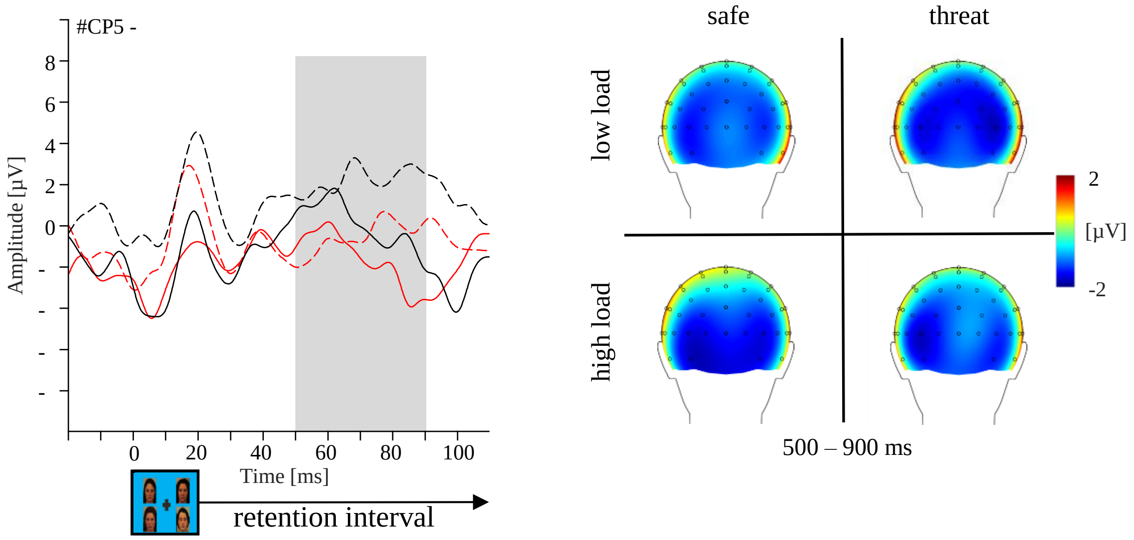


***Figure 10*. VWM: load and context effects for N170 and CDA**. Illustration of the load and context effect during change detection. (A) Grand averaged N170 plotted as a function of memory Load (low [1 face] vs. high [2 faces]) and Context (threat vs. safety) for an exemplary parieto-occipital sensor and a topographical difference map (high – low load) displaying the averaged time interval of peak detection (120 – 190 ms) plotted on a back view of a model head. (B) CDA (contralateral minus ipsilateral), recorded at the CP5-CP6 sites, time-locked to the onset of the memory array, plotted as a function of set size (one face (low load) vs. two faces (high load) and as a function of context (threat vs. safety). Waveforms were filtered with a high cutoff filter of 10 Hz for visual inspection only. Top views of topographical maps for each experimental condition (one vs. two faces and threat vs. safety contexts) recorded during a 500–900-ms interval following the onset of the memory array.

3.5 Discussion

Summary of main findings. This preregistered study examined effects of social threat/safety learning – through verbal instructions or observing others – on face perception and memory in individuals with adverse childhood experiences. While participants experienced contextual threat of receiving shocks or safety, we focused on two memory modalities (item/source and visual working memory). As confirmed by self-reported ratings, the threat context was perceived as more arousing, threatening, and unpleasant relative to the safety condition. Interestingly, social threat induction was equally effective for both verbal and observational learning, and this was also evident for electrocortical processing. At the behavioral level, contextual safety seemed to facilitate face recognition in the item/source memory task, especially for the more socially anxious participants. In contrast, visual working memory was not affected by context conditions. However, the more socially anxious participants appear to perform better on threat than on safety and have a lower false alarm rate for detecting changes in the visual field. In addition, as expected, visual working memory was impaired in the more demanding task (high task load with four faces).

Regarding electrocortical activity, threat-selective processing patterns were observed during memory encoding in the item/source memory task. Confirming an old-new recognition effect, previously seen faces (old) revealed differential processing relative to unknown (new) faces. Interestingly, although behavioral recognition performance was overall poor, threat-selective face processing was observed during recognition (late parieto-occipital positivity for threat compared to safe and new faces; Schellhaas et al., 2020). Regarding visual working memory, task load reduced processing as revealed by N170 and CDA (late time window) in the high load condition. Consistent with the behavioral results, the threat context gained more access to working memory and decreased its capacity, as evidenced by an increased CDA for the threat compared to the safe context. Overall, the present data suggest that regardless of the severity of childhood trauma, contextual threat alters face perception and reduces face recognition and storage, especially when cognitive demands are high. However, individuals with high social and state anxiety appear to benefit from the threatening contextual conditions and show improved detection of visual changes.

Item and source memory: Safety-enhanced face recognition but poor source memory. For the item/source memory task, participants were asked to indicate whether they have seen a face (i.e., the item) before and, if so, in what context (i.e., threat or safety source). On the electrocortical level, differential face processing was observed for faces presented during threat relative to safety in the encoding phase at two processing stages (parieto-occipital 250-400ms and fronto-central 540-800ms). In addition, we replicated an old/new recognition effect showing enhanced processing for correctly recognized old faces (parieto-occipital 660-800ms, and central 750-900ms), which also varied for threat compared to safe and new faces (parieto-occipital 700-900ms). Interestingly, this neural differentiation between threat- and safety-associated faces did not result in improved recognition memory, i.e., participants were unable to consciously identify the context in which they had previously encountered a face. This dissociation of neural measures and behavioral recognition performance is a direct replication of a previous study in a healthy sample (Schellhaas et al., 2020), and goes along with studies showing that neural processing is more sensitive to subtle context differentiation than behavioral performance (Sessa et al., 2011; Luck et al., 1996). It is therefore suggested that an observed dissociation between ERPs and behavior is not due to the absence of processing modulations underlying a cognitive task (Wilkinson & Halligan, 2004), but may vary with subtle neural processing not resulting in overt behavioral changes.

Building upon previous research (Ventura-Bort et al., 2016b; Kensinger, 2009), we predicted that an arousing context will enhance face memory. This hypothesis is based on the Arousal Biased Competition model (ABC; Mather & Sutherland, 2011), which assumes that processing of arousing and therefore salient stimuli is enhanced, whereas less salient stimuli must compete more vigorously for attentional resources. In contrast, our results showed a tendency for better recognition of faces encoded during safety, while the context was not remembered at all. Several alternative explanations may account for these diverging findings. Bisby et al. (2018) suggest that the presence of a negative (context) element weakens the item-context association while leading to an enhancement in item recognition. However, this presupposes a temporal or conceptual link between face and context (item-source binding; Dunsmoor et al., 2015), which was not the case in the current study because face-context associations were not reinforced. Alternatively, attentional resources may have been allocated exclusively to the threatening context during encoding, resulting in enhanced processing of threat versus safety/new during both encoding and recognition, but impaired recognition of faces from the threatening context. Finally, new information and especially faces need time to consolidate (McGaugh, 2000). For instance, memory for faces is strongly depending on familiarity with a person (Burton & Jenkins, 2011) and neutral looking unknown faces are per se difficult to recognize (Arnold, González, Schellhaas, & Bublatzky, 2021). Here, explicit learning instructions may enhance the strength of face-context association, and longer consolidation periods between encoding and recognition phase seem pertinent to boost item and source memory (e.g., a gap of one hour, day or even week; Ventura-Bort et al., 2016b).

Visual working memory: Reduced capacity with high load and threat. The visual working memory task required detection of changes in a visual scene displaying either two or four face identities (low and high task load). In accordance with our hypotheses, change detection performance was impaired under high task load, suggesting limited cognitive resources (Sessa et al., 2011). This effect was also found in electrocortical processing, which showed increased contralateral delay activity (CDA) and N170 for the high- compared to low-load condition, reflecting changes in working memory capacity (Morgan et al., 2008). Regarding the impact of threat, the contextual CDA modulation confirmed previous findings showing a reduced capacity for threat (Stout et al., 2013). Thus, threatening information plays a key role in the modulation of the CDA showing an increasing amplitude the more relevant the threatening information becomes (Ward et al., 2020).

Again, result patterns for perceptual processing and overt behavioral performance do not match, which may indicate low behavioral relevance of contextual threat in the present study. In contrast to that notion, more socially anxious participants were better in detecting visual changes within a threatening context (lower false alarm rates, higher k and d’ parameter). This finding is in line with the Attentional Control Theory (Eysenck et al., 2007), which assumes that arousal consumes working memory resources and thereby reduces the ability to actively inhibit distracting information. This claim is supported by several studies showing that high state, trait and social anxiety is associated with attentional distraction by task irrelevant stimuli (Eysenck & Derakshan, 2011; Moriya & Sugiura, 2012). Thus, anxiety interferes with working memory processes (inhibition shifting, updating) and persons with high trait anxiety levels perform worse on demanding tasks, or alternatively, must put in compensatory efforts that lead to a good task performance with a trade-off in efficiency (slower reaction times; Edwards et al., 2015). In the present change detection task (with no distracting elements), contextual threat could have led to heightened attention and alertness in highly anxious participants, thereby boosting memory performance during threat (Robinson et al., 2013).

Social threat and safety learning: Verbal and observed information is similarly effective. Contextual threat or safety was learned through verbal instructions or observing a model receiving electrical shocks. Similar to previous research (Olsson & Phelps, 2007; Bublatzky et al., 2020), such social means of learning were highly effective in establishing a threatening context, which was perceived as more threatening, unpleasant and arousing, and elicited threat-selective electrocortical processing compared to the safety context (Schellhaas et al., 2020; Arnold et al., 2021). Interestingly, however, these threat effects were independent of social learning type. Specifically, no differences were found between learning through verbal instructions and observing others neither for ratings, behavior or electrocortical measures. This finding is in line with recent research suggesting that both verbal and observational threat learning show indistinguishable behavioral effects (comparable to direct threat learning; Olsson & Phelps, 2007). However, regarding the underlying neural network partly different regions have been suggested to be involved in verbal and observational learning (e.g., ACC, anterior insula; Lindström et al., 2018). Exploring these differences and how to counteract maladaptive social threat learning, for example, when it leads to pathological anxiety, is an important goal for future research. In particular, supporting correct recall of threatening and/or safe situations may prove helpful in overcoming psychopathological anxious arousal.

Adverse childhood experiences and inter-individual differences in anxiety. The impact of adverse childhood experiences (ACE) can be diverse and different types and timings can have distinct effects on neural development, behavior, and psychopathology (Sheridan & McLaughlin, 2014; Herzog & Schmahl, 2018). In this study, we followed a trans-diagnostic approach with the childhood trauma questionnaire (CTQ) as a dimensional measure. This resulted in a heterogeneous sample of participants who experienced varying degrees of childhood abuse and neglect, often simultaneously, ranging from milder CTQ scores to moderate and severe effects of childhood maltreatment. In this highly ecologically valid sample, we found intact memory processing despite the presence of sometimes severe ACE and comorbid psychopathology. Contrary to our hypotheses, we did not find overall altered or generalized threat processing associated with ACE, nor did we find changes in item/source memory or visual working memory, even independent of the cognitive load acting on the memory system (Goodman et al., 2019). Instead, we found that contextual safety improved recall of unknown neutral faces independent of recall of the context and severity or type of ACE. Exploratory analyses of inter-individual differences in social and state anxiety showed that this benefit in face recognition increased in more anxious participants.

Regarding subtypes of ACE, exploratory analyses showed that sexual abuse and emotional neglect were associated with improved recognition of faces from a threatening context. In addition, sexual and emotional abuse were associated with improved working memory during threat, possibly due to heightened vigilance in stressful situations. In contrast, emotional neglect scores were associated with better working memory under safe conditions (Robinson et al., 2013). This may reflect greater recruitment of cognitive resources in order to maintain task performance during low-stress tasks, resulting in equal or improved performance compared to individuals with low anxiety, but at the expense of highly effortful task performance (Moriya & Sugiura, 2013). Thus, when task demands are increased by distracting information, inhibition and filtering of these distractions are impaired and performance declines, especially when the distractions are associated with threat (Stout et al., 2013). While the tasks at hand were not designed to compare varying degrees of distraction and inhibition as part of the memory processes, however, future studies could be designed to include both threatening and neutral distractors into the memory task to investigate compensatory effects (Ward et al., 2020; MacNamara et al., 2011).

Limitations and future directions. First, because the task were demanding due to their difficulty (item and source memory) and effort (visual working memory), motivation plays an important role. Participants spent approximately 30 minutes on each task, with a short break in between, resulting in a long individual task and total study time. Motivation is related to cognitive control and load, and a lack of it can lead to decreased performance. To increase motivation, performance-based incentives or a time limit for responses could be introduced, especially in the visual working memory task, which in turn could lead to a reduced influence of the motivational effect.

Second, our ACE sample included less severe cases and therefore more high-functional individuals (in terms of memory functioning). Besides self-selection, our exclusion criteria were rather strict (e.g., no medication except SSRIs and SNRIs, no drug intake or comorbid addiction). Because the more severe cases of early traumatization are more likely to develop severe psychiatric disorders and comorbidities leading to the use of psychotropic medications, the pool of severely traumatized participants in the current study is limited. This in turn could have led to comparable memory performance with regard to healthy participants (Schellhaas et al., 2020). Future studies may consider more liberal inclusion criteria.

Third, the results show a dissociation between perceptual/attentional processes and behavioral measures. Future research should address whether this dissociation is reflected in psychophysiological response priming, for instance, whether participants show autonomic arousal to faces encountered in a threatening context but not retrieved from memory. This is also of clinical relevance, as identifying unconscious arousal triggers is a key element of many treatments (e.g., for PTSD; Cahill & Foa, 2007). Moreover, in the age of a highly socially networked world (e.g., social media), the importance of socially mediated threat triggers is increasing in general and in psychopathology. Finally, both type and timing of adverse events during childhood play a critical role and have an impact on the brain. For instance, sensitive periods and specific ACE-subtypes influence the development of neurobiological alterations differentially (e.g., cortical thickness; Herzog & Schmahl, 2018), leading to differential effects of memory processes in adulthood. Therefore, more sensitive measures of the type and timing of ACE should be included in future studies.

Conclusions. This study shows that contextual threat and safety lead to processing differences in face perception, recognition, and change detection in adult individuals with adverse childhood experiences. While overall intact social threat and safety learning was found, threat-selective face processing was observed in an item/source memory task and the threat context required more processing resources in a visual working memory task. Regarding behavioral performance, generally safety seemed to facilitate face recognition in low anxious individuals with low/moderate levels of ACE. In contrast, highly anxious and severely maltreated individuals showed better task performance in aversive anticipatory contexts. More research is needed to examine the psychophysiological processes involved in functional and dysfunctional memory systems, and their relevance as vulnerability factors for stress-related disorders.

3.6 Supplemental Material

Supplement 1: Hierarchical multinomial processing tree modeling (MPT)

Hierarchical multinomial processing tree (MPT) modelling was used to disentangle item recognition, source recognition, and guessing biases for the item/source memory task (Batchelder & Riefer, 1999; Bayen et al., 1996). Discrete observable states are used to draw conclusions of non-observable underlying cognitive processes. MPT models assume that the observed categorical responses (answer frequencies, e.g., for classifying a face as old) follow a multinomial distribution. The model parameters represent transition probabilities between latent cognitive states (e.g., knowing or guessing; Matzke et al., 2015; for detailed reviews, see Batchelder & Riefer, 1999; Erdfelder et al., 2009; Arnold, Bayen, & Smith, 2015; see Fig. 2). The MPT model we used is the two-high-threshold model of source monitoring (2HTSM; Bayen et al., 1996). Faces are presented with one of two sources, here within either a threatening or safe context, or faces are new. Therefore, the model consists of three decision trees: (1) faces presented in the threat condition, (2) faces presented in the safety condition, and (3) new faces (i.e., that were not presented during the encoding phase; Figure 7).

Participants correctly recognize a face as old or new with probability D_T (from threat context), D_S (from safe context), and D_N (new). With probability d_T participants correctly recognize the threatening source, with probability d_S they correctly recognize the safe source. If they do not remember the source of the face (probabilities $1-d_T$ and $1-d_S$ for threat and safety, respectively), they must guess. With probability g, participants guess the threatening source and with probability 1-g the safe source. If participants do not recognize that a face was old or new (probability $1-D_{T/S/N}$), they guess with probability b that a face was old and therefore with probability 1-b that a face was new. The overall probability of a certain response category then results from the sum of the probabilities of the individual branches. For example, the probability for answering "Threat" when an item stems from a threatening source is $p(\text{"Threat"}) = D_T*d_T + D_T*(1-d_T)*g+(1-D_T)*b*g$.

Hierarchical Bayesian extensions for MPT assume that the individual parameters follow a continuous hierarchical distribution, accounting for differences between participants (e.g., Heck et al., 2018; Klauer, 2010; Smith & Batchelder, 2010) and allow including individual questionnaire scores in the MPT model (e.g., Arnold, Bayen, & Böhm, 2015; Schaper et al., 2019). Therefore, model parameters were estimated on group level as well as on individual level. The model in this form is locally not identifiable, therefore we restricted the recognition parameters for safe and new faces to be equal ($D_S = D_N$, Keefe et al., 2002). As a between-factor, Learning (observation vs. instruction) was included, separate parameter estimates were calculated for each of the three decision trees.

Supplement 2: Additional sample characteristics regarding psychopathology

The most common disorders were major depressive disorder, recurrent episode (lifetime 54.4% and current 19%) and post-traumatic stress disorder (lifetime 43.9% and current 26.3%), with most of the participants having at least one comorbidity, mostly substance (alcohol) abuse disorder in the past (14.3%) and social anxiety disorder in the past (20%) and at the time of recruitment (16.4%; see Supplementary Table S1).

Table S1. *Overview of diagnostics (according to DSM-IV and DSM-5).*

N = 59 Age = 33 (±11) 54 female	Current (%)	Past (%)
Bipolar and Related Disorders		
Bipolar I Disorder	–	1.7
Depressive Disorders		
Major Depressive disorder, single episode	–	17.9
Major Depressive Disorder, Recurrent Episode	**19**	**54.4**
Persistent Depressive Disorder (Current = 2 Years)	8.8	10.7
Substance Use Disorder (Current = 12 Months)		
Alcohol Use Disorder	**1.7**	**14.3**
Sedative, Hypnotic, Or Anxiolytic Use Disorder	1.7	–
Cannabis Use Disorder	–	7.1
Amphetamine-Type Substance Use Disorder	–	3.6
Cocaine Use Disorder	–	1.8
Opioid Use Disorder	–	1.8
Other Hallucinogen Use Disorder	–	1.8
Inhalant Use Disorder	–	1.8
Anxiety And Trauma- And Stressor-Related Disorders		
Panic Disorder	7.3	7.3
Agoraphobia (Current = 6 Months)	1.8	9.1
Social Anxiety Disorder (Current = 6 Months)	**16.4**	**20.0**
Generalized Anxiety Disorder (Current = 6 Months)	7.3	9.1
Obsessive-Compulsive Disorder (Current = 1 Month)	5.5	7.3
Post-Traumatic Stress Disorder	**26.3**	**43.9**
Attention-Deficit/Hyperactivity Disorder (Current = 6 Months)		
Combined Presentation	7.0	8.9
Predominantly Inattentive Presentation	7.0	7.1
Predominantly Hyperactive/Impulsive Presentation	1.8	1.9
Other Disorders		
Substance/Medication-Induced Obsessive-Compulsive and Related Disorder	5.4	9.1
Premenstrual Dysphoric Disorder	6.4	5.9
Specific Phobia		
Fear Of Blood	1.8	1.8
Fear Of Injections and Transfusions	1.8	1.8
Situational Fear	1.8	1.8
Body Dysmorphic Disorder	–	1.7
Anorexia Nervosa	–	14.3
Bulimia Nervosa	1.8	10.7
Binge-Eating Disorder	1.7	3.6
Somatic-Symptom Disorder	1.7	5.3
Schizophrenia Spectrum and Other Psychotic Disorders		
Brief Psychotic Disorder	–	3.5

Supplement 3: Extreme group analyses (based on SPIN, STAI-S and CTQ)

Exploratory extreme group analyses were based on the distribution of the SPIN, STAI-S and CTQ questionnaires (i.e., low, moderate, and high-scorers; see Supplementary Tables S2). Cut-off values for low groups were SPIN ≤ 15, STAI-S ≤ 36, CTQ ≤ 57, for medium groups 28 ≥ SPIN > 15, 39 ≥ STAI-S > 36, 79 ≥ CTQ > 57, and for high groups SPIN > 28, STAI-S > 39, CTQ > 79.

For SPIN and STAI-S groups, there were no differences regarding age and other questionnaire measures (Age: $F < 1$, SPIN: $F(2,60) = 2.30$, $p = .11$, and CTQTOT: $F(2,60) = 3.10$, $p = .05$ for STAI-S groups; Age: $F(2,60) = 1.06$, $p = .35$; STAI-S: $F < 1$, and CTQ_{TOT}: $F(2,60) = 2.31$, $p = .11$ for SPIN groups). Post-hoc test reveal significant differences between the low, moderate, and high sub-groups for SPIN, $F(2,61) = 164.84$, $p < .001$, and for STAI-S: $F(2,60) = 46.16$, $p < .001$. For CTQTOT, groups did not differ in anxiety, SPIN: $F(2,59) = 2.06$, $p = .14$, STAI-S: $F(2,59) = 1.33$, $p = .27$, but they differed regarding age distribution, Age: $F(2,59) = 3.70$, $p < .05$, with post hoc tests revealing differences between low-high and medium-high groups.

Inter-individual differences based on extreme groups in item/source memory. Exploratory extreme group analyses revealed overall higher item recognition for the low social anxiety group compared to the moderate group, SPIN-Group $F(2,61) = 4.03$, $p < .05$, $\eta_p^2 = .12$ [.01, .24] (HR-FAR: $M_{\text{low}} = .54$ (SD = .03), $M_{\text{med}} = .43$ (SD = .03), $M_{\text{high}} = .47$(SD = .03)).

Inter-individual differences based on extreme groups in visual working memory. Exploratory extreme group analyses showed a significant interaction for SPIN Group × Learning, $F(2,56) = 3.26$, $p < .05$, $\eta_p^2 = .10$ [.00, .20 for the HR. Post-hoc tests revealed that the hit rate was higher for the instructional compared to the observational learning group in the high SPIN group only.

For the FAR, a significant effect based on exploratory SPIN grouping was found, showing that the FAR was higher for the medium SPIN group compared to both, the low and high SPIN Group, *F*(2,56) = 3.59, $p < .05$, $\eta_p^2 = .11$ [.01, .23]. Additionally, for the FAR a significant SPIN Group × Context interaction was found, *F*(2,56) = 3.82, $p < .05$, $\eta_p^2 = .12$ [.01, .24], with post-hoc tests revealing a lower FAR for the threat compared to safe context only in the high SPIN group. For the groups based on STAI-S, a significant three-way interaction effect for FAR between Context × Load × STAI-S Group was found, indicating that the context differs in the high load condition for the medium STAI-S Group, and the low and high STAI-S group differ in the low load condition during threat, *F*(2,55) = 4.12, $p < .05$, $\eta_p^2 = .14$ [.01, .25].

When SPIN groups were included as a between factor, a Context × Learning interaction effect for d‘ emerged, *F*(1,55) = 4.46, $p < .05$, $\eta_p^2 = .08$ [.02, .26], showing higher change detection (d‘) for threat in the observation group only.

Similarly to the main analyses, for the exploratory group analyses, the capacity measure (k) revealed a significant Load × SPIN group interaction, with higher k for the high compared to the medium SPIN Group only in the high load condition, *F*(2,56) = 3.32, $p < .05$, $\eta_p^2 = .12$ [.00, .18]. A significant Context × Learning interaction showed higher k values for the threatening context in the observation group only, *F*(2,56) = 4.42, $p < .05$, $\eta_p^2 = .07$ [.02, .26]. For capacity and change detection (d‘ and k), a significant Context × Learning interaction showed higher values for d‘ and k for the threat compared to the safe context in the observation group only, *F*(2,54) = 5.80, $p < .05$, $\eta_p^2 = .10$ [.03, .30] and *F*(2,55) = 4.19, $p < .05$, $\eta_p^2 = .07$ [.01, .26], respectively.

Inter-individual differences based on extreme groups for ERPs in item/source memory. Old/New recognition: Exploratory extreme group analyses revealed that the Old/New effect (+correct source identification) varied as a function of CTQ_{TOT} Group. Post-hoc tests indicate an overall enhanced processing of both old and new faces for the low compared to the high CTQ_{TOT} Group, *F*(2,56) = 3.31, $p < .05$, $\eta_p^2 = .11$ [.00, .23] (M_{low} = 2.27 (SD = .39), M_{med} = 1.31 (SD = .39), M_{high} = .93 (SD = .37)). No effects were found for STAI-S or SPIN groups.

Source recognition: For the exploratory extreme group analyses, a significant effect of SPIN-Group was found for the recognition phase (700-900ms), $F(2,57) = 3.64$, $p < .05$, $\eta_p^2 =$ .11 [.01, .23] ($M_{low} = 4.32$ (SD = .64), $M_{med} = 2.11$ (SD = .69), $M_{high} = 2.20$ (SD = .69)), showing generally enhanced amplitudes for the group with low levels of social anxiety compared to the moderate and high anxious groups. No effects were found for STAI-S or CTQ_{TOT}.

Tables S2. *Summary of the behavioral performance in the change detection task (VWM) based on exploratory extreme groups.*

(A)

		Sex (♀/♂)	Age	SPIN	STAI-S	CTQ_{TOT}
SPIN	Low	20/3	30.34 (10.10)	10.13 (3.84)	38.59 (1.53)	65 (21.40)
	Medium	19/1	32.20 (11.70)	21.80 (4.06)	37.80 (2.82)	72.21 (24.36)
	High	20/1	35 (11.41)	39.57 (7.55)	37.38 (5.14)	82.57 (24.02)
STAI-S	Low	20/1	32.81 (10.75)	28.71 (13.13)	32.52 (5.10)	82.67 (25.19)
	Medium	15/1	29.19 (9.99)	21.06 (13.31)	38.19 (.83)	69.81 (21.71)
	High	23/3	34.31 (12.01)	21.35 (12.86)	42.15 (2.62)	67.13 (23.26)
CTQ	Low	18/2	29.50 (11.86)	20.30 (10.80)	38.85 (3.80)	47.55 (6.50)
	Medium	19/1	29.75 (9.77)	21.65 (13.10)	38.35 (7.13)	68.60 (6.57)
	High	20/2	37.09 (9.36)	28.09 (15.53)	36.29 (4.63)	100.59 (12.84)

(B)

		HR – observation					HR – instruction			
		low load		high load			low load		high load	
		threat	safe	threat	safe		threat	safe	threat	safe
SPIN	Low (N = 12)	.82 (.21)	.82 (.18)	.52 (.11)	.51 (.11)	Low (N = 10)	.81 (.19)	.82 (.18)	.51 (.11)	.52 (.11)
	Medium (N = 6)	.83 (.16)	.79 (.19)	.57 (.07)	.51 (.05)	Medium (N = 13)	.71 (.19)	.72 (.19)	.51 (.10)	.54 (.12)
	High (N = 12)	.72 (.25)	.73 (.24)	.47 (.13)	.44 (.12)	High (N = 9)	.87 (.12)	.88 (.10)	.59 (.09)	.57 (.11)
STAI-S	Low (N = 11)	.69 (.21)	.72 (.19)	.50 (.09)	.48 (.09)	Low (N = 9)	.82 (.12)	.84 (.10)	.54 (.07)	.54 (.08)
	Medium (N = 9)	.81 (.24)	.75 (.25)	.48 (.16)	.47 (.14)	Medium (N = 6)	.75 (.24)	.78 (.21)	.54 (.07)	.50 (.11)
	High (N = 9)	.85 (.20)	.85 (.16)	.54 (.11)	.49 (.11)	High (N = 17)	.78 (.19)	.77 (.20)	.53 (.13)	.56 (.12)
CTQ	Low (N = 7)	.85 (.17)	.88 (.07)	.53 (.09)	.53 (.06)	Low (N = 13)	.76 (.22)	.75 (.23)	.53 (.11)	.56 (.13)
	Medium (N = 8)	.80 (.22)	.77 (.22)	.51 (.10)	.50 (.11)	Medium (N = 11)	.86 (.16)	.85 (.14)	.54 (.10)	.52 (.11)
	High (N = 14)	.73 (.25)	.72 (.24)	.50 (.14)	.44 (.12)	High (N = 7)	.74 (.12)	.81 (.10)	.51 (.12)	.54 (.08)

(C)

		FAR – observation					FAR – instruction			
		low load		high load			low load		high load	
		threat	safe	threat	safe		threat	safe	threat	safe
SPIN	Low (N = 12)	.17 (.08)	.17 (.10)	.18 (.09)	.21 (.10)	Low (N = 10)	.26 (.12)	.24 (.11)	.24 (.17)	.23 (.14)
	Medium (N = 6)	.22 (.21)	.22 (.18)	.32 (.20)	.31 (.23)	Medium (N = 13)	.34 (.16)	.32 (.16)	.38 (.22)	.32 (.20)
	High (N = 12)	.18 (.15)	.24 (.14)	.16 (.09)	.17 (.09)	High (N = 9)	.22 (.12)	.22 (.18)	.18 (.10)	.24 (.11)
STAI-S	Low (N = 11)	.14 (.07)	.19 (.12)	.20 (.11)	.20 (.13)	Low (N = 9)	.19 (.13)	.19 (.16)	.16 (.12)	.19 (.11)
	Medium (N = 9)	.17 (.19)	.19 (.15)	.22 (.16)	.20 (.16)	Medium (N = 6)	.28 (.16)	.27 (.17)	.38 (.22)	.29 (.17)
	High (N = 9)	.24 (.14)	.23 (.16)	.18 (.13)	.24 (.14)	High (N = 17)	.33 (.12)	.31 (.14)	.31 (.19)	.30 (.17)
CTQ	Low (N = 7)	.21 (.14)	.22 (.17)	.19 (.08)	.20 (.08)	Low (N = 13)	.27 (.17)	.26 (.18)	.31 (.26)	.29 (.20)
	Medium (N = 8)	.21 (.18)	.23 (.15)	.24 (.20)	.27 (.20)	Medium (N = 11)	.29 (.13)	.26 (.14)	.25 (.12)	.25 (.13)
	High (N = 14)	.16 (.11)	.19 (.11)	.20 (.10)	.19 (.12)	High (N = 7)	.26 (.09)	.26 (.15)	.26 (.16)	.23 (.12)

(D)

		d' – observation					d' – instruction			
		low load		high load			low load		high load	
		threat	safe	threat	safe		threat	safe	threat	safe
SPIN	Low (N = 12)	2.21 (1.01)	2.17 (.92)	1.03 (.41)	.88 (.39)	Low (N = 10)	1.76 (1.04)	1.83 (.97)	.81 (.59)	.88 (.48)
	Medium (N = 6)	1.97 (1.06)	1.83 (1.24)	.73 (.54)	.73 (.82)	Medium (N = 13)	1.18 (1.04)	1.19 (.97)	.43 (.77)	.72 (.62)
	High (N = 12)	1.85 (1.01)	1.59 (.91)	.98 (.56)	.85 (.41)	High (N = 9)	2.10 (.80)	2.20 (1.00)	1.21 (.38)	.95 (.39)
STAI-S	Low (N = 11)	1.76 (.97)	1.69 (1.10)	.92 (.43)	.91 (.56)	Low (N = 9)	2.00 (.99)	2.11 (1.02)	1.22 (.58)	1.07 (.44)
	Medium (N = 9)	2.21 (1.12)	1.84 (1.14)	.76 (.54)	.87 (.45)	Medium (N = 6)	1.58 (1.33)	1.65 (1.17)	.50 (65)	.67 (.51)
	High (N = 9)	2.08 (.99)	2.02 (.87)	1.17 (.52)	.77 (.50)	High (N = 17)	1.43 (.94)	1.40 (.97)	.62 (.67)	.77 (.53)
CTQ	Low (N = 7)	2.22 (.99)	2.16 (.89)	1.02 (.36)	.94 (.42)	Low (N = 13)	1.57 (1.22)	1.57 (1.24)	.73 (.88)	.86 (.61)
	Medium (N = 8)	1.97 (1.19)	1.76 (1.21)	.90 (.64)	.73 (.60)	Medium (N = 11)	1.94 (1.01)	1.83 (.87)	.85 (.56)	.83 (.50)
	High (N = 14)	1.90 (.97)	1.72 (.93)	.89 (.47)	.84 (.50)	High (N = 7)	1.35 (.59)	1.72 (.97)	.75 (.58)	.89 (.37)

(E)

		k – observation					k – instruction			
		low load		high load			low load		high load	
		threat	safe	threat	safe		threat	safe	threat	safe
SPIN	Low (N = 12)	1.30 (.50)	1.31 (.47)	1.36 (.47)	1.20 (.48)	Low (N = 10)	1.10 (.56)	1.15 (.54)	1.07 (.77)	1.18 (.62)
	Medium (N = 6)	1.23 (.61)	1.13 (.73)	1.01 (.70)	.81 (.78)	Medium (N = 13)	.75 (.56)	.80 (.61)	.53 (.88)	.88 (.58)
	High (N = 12)	1.08 (.54)	.99 (.48)	1.21 (.56)	1.06 (.48)	High (N = 9)	1.29 (.41)	1.31 (.46)	1.65 (.45)	1.34 (.49)
STAI-S	Low (N = 11)	1.09 (.51)	1.06 (.50)	1.20 (.46)	1.13 (.56)	Low (N = 9)	1.26 (.47)	1.30 (.48)	1.51 (.61)	1.39 (.39)
	Medium (N = 9)	1.27 (.60)	1.1.2 (.68)	1.04 (.74)	1.10 (.62)	Medium (N = 6)	.95 (.70)	1.03 (.64)	.65 (.82)	.85 (.53)
	High (N = 9)	1.22 (.53)	1.24 (.48)	1.44 (.44)	1.00 (.54)	High (N = 17)	.90 (.54)	.93 (.59)	.87 (.91)	1.04 (.65)
CTQ	Low (N = 7)	1.29 (.47)	1.32 (.44)	1.38 (.40)	1.30 (.51)	Low (N = 13)	.97 (.70)	.97 (.75)	.89 (1.07)	1.08 (.70)
	Medium (N = 8)	1.17 (.64)	1.07 (.67)	1.08 (.66)	.92 (.67)	Medium (N = 11)	1.14 (.48)	1.16 (.43)	1.18 (.74)	1.09 (.53)
	High (N = 14)	1.14 (.52)	1.06 (.50	1.20 (.56)	1.02 (.50)	High (N = 7)	.96 (.35)	1.11 (.46)	1.01 (.74)	1.25 (.47)

(F)

		RT – observation					RT – instruction			
		low load		high load			low load		high load	
		threat	safe	threat	safe		threat	safe	threat	safe
SPIN	Low (N = 11)	675.98 (253.39)	703.56 (215.42)	1004.53 (390.76)	929.06 (344.24)	Low (N = 10)	690.07 (589.80)	669.44 (388.44)	780.89 (299.00)	861.14 (854.42)
	Medium (N = 6)	709.77 (140.22)	682.14 (142.24)	855.64 (191.09)	848.30 (96.83)	Medium (N = 13)	1220.09 (790.84)	1225.13 (785.70)	1650.94 (1133.20)	1726.26 (1289.34)
	High (N = 12)	886.23 (470.09)	963.36 (578.91)	1188.97 (1033.73)	1180.65 (785.71)	High (N = 9)	1275.39 (1582.18)	775.46 (368.33)	992.99 (484.72)	877.08 (291.69)
STAI-S	Low (N = 11)	847.92 (279.40)	837.11 (282.19)	996.23 (302.91)	1020.74 (376.25)	Low (N = 9)	1208.10 (1606.91)	792.87 (303.51)	906.43 (306.17)	844.14 (271.66)
	Medium (N = 9)	693.09 (332.49)	747.06 (349.55)	889.62 (418.26)	948.12 (589.50)	Medium (N = 6)	1233.70 (767.22)	1147.26 (633.18)	1315.87 (720.63)	1628.64 (1212.35)
	High (N = 9)	782.04 (477.72)	843.10 (644.42)	1345.14 (1228.38)	1096.07 (784.35)	High (N = 17)	954.70 (736.68)	931.97 (751.20)	1303.23 (1075.17)	1294.76 (1203.24)
CTQ	Low (N = 7)	839.64 (520.47)	932.05 (692.98)	1392.34 (1279.43)	1171.97 (811.63)	Low (N = 13)	1054.07 (865.89)	979.47 (706.08)	1171.97 (910.78)	1343.97 (1239.18)
	Medium (N = 8)	748.13 (273.32)	710.97 (264.72)	1045.83 (494.64)	914.78 (352.08)	Medium (N = 11)	714.24 (404.00)	658.63 (336.52)	1098.49 (886.06)	841.38 (371.75)
	High (N = 14)	747.62 (311.23)	798.59 (303.86)	881.31 (302.81)	996.81 (520.52)	High (N = 7)	1625.82 (1755.74)	1061.65 (518.74)	1199.95 (740.37)	1259.70 (901.22)

Note. Values represent mean (M) and standard deviation (SD): M (SD); (A) characteristics of the groups based on the questionnaires: SPIN = social phobia inventory; STAI-S = State/Trait anxiety inventory – state; CTQ = childhood trauma questionnaire; (B) – (F) performance in the different behavioral measures of VWM: HR = hit rate; FAR = false alarm rate; HR – FAR = recognition rate; d‘ = change detection ability; k = capacity index; RT = reaction time (in ms), separately for the low/high load and threat/safe context conditions.

Supplement 4: Supplementary Analyses

Participants had longer reaction times during high load, irrespective of context or learning, Load *F*(1,60) = 14.16, $p < .001$, $\eta_p^2 = .20$ [.06, .33], Context *F*(1,60) < 1, $p = .55$, Context × Load F(1,60) < 1, $p = .57$, Context × Learning: *F* < 1, Load × Learning: *F* < 1, Learning: *F*(1,60) < 1, $p = 0.24$. A significant covariation occurred for Load × emotional abuse, *F*(1,54) = 4.61, $p < .05$, $\eta_p^2 = .08$ [.00, .19], and Load × emotional neglect, *F*(1,54) = 7.08, $p < .05$, $\eta_p^2 = .10$ [.02, .25], indicating slower reaction times in the high load for highly emotionally abused participants and faster reactions of highly emotionally neglected individuals. For RT, STAI-S groups varied with Load, showing higher RT for the high load condition compared to the low load condition only in the high STAI-S Group, *F*(2,53) = 3.41, $p < .05$, $\eta_p^2 = .11$ [.00, .24]. There were no significant effects based on CTQ_{TOT} Groups.

A test of sequence/order of memory modality did not reveal any effects on the conventional performance measures of the ISM task. For the VWM task, a 3-way interaction effect for the FAR was found, *F*(1,60) = 4.34, $p < .05$, $\eta_p^2 = .07$ [.00, .20]. Post-hoc tests show that if the ISM task was completed before the VWM task, the FAR for the threat context was lower compared to the FAR in the threat context in the low load condition, when the VWM was first. This could be a small hint of a decline of threat/aversive interference with task performance.

Covariation between electrocortical and behavioral data. For explorative purposes, performance measures were included in the ERP analyses (HR-FAR for old/new effect, HR-FAR, and ACSIM [threat/safety only] for context comparisons (700-900ms). The only significant covariation emerged for Context × ACSIM threat, showing reduced differences between threat and safety with higher ACSIM threat, *F*(1,57) = 4.39, $p < .05$, $\eta_p^2 = .07$ [.00, .19].

Additionally, ERP measures were included as covariates (Old-New, Threat-Safe, Threat-New, Safe-New, Old, New, Threat, Safe) into the MPT model. No meaningful relation with the MPT parameters occurred.

Supplement 5: Stimulus material

Neutral face pictures of in total 150 actors (half females; 90 were chosen for the item/source memory task and 60 for the visual working memory task) were selected from the Karolinska Directed Emotional Faces (Lundqvist et al., 1998), the Radboud (Langner et al., 2010), and the NimStim database (Tottenham et al., 2009).

Item source memory

Female

010_y_f_n_a
020_y_f_n_a
028_y_f_n_a
048_y_f_n_a
069_y_f_n_a
071_y_f_n_a
090_y_f_n_a
09F_NE_C
115_y_f_n_a
134_y_f_n_a
163_y_f_n_a
af01nes
af02nes
af04nes
af05nes
af08nes
af09nes
af11nes
af12nes
af13nes
af21nes
af23nes
af25nes
af26nes
af27nes
af29nes
af32nes
af33nes
af34nes
F8neut_c_st
Rafd090_01_Caucasian_female_neutral_frontal
Rafd090_02_Caucasian_female_neutral_frontal
Rafd090_04_Caucasian_female_neutral_frontal
Rafd090_08_Caucasian_female_neutral_frontal
Rafd090_14_Caucasian_female_neutral_frontal
Rafd090_22_Caucasian_female_neutral_frontal
Rafd090_26_Caucasian_female_neutral_frontal
Rafd090_27_Caucasian_female_neutral_frontal
Rafd090_31_Caucasian_female_neutral_frontal
Rafd090_32_Caucasian_female_neutral_frontal
Rafd090_37_Caucasian_female_neutral_frontal
Rafd090_56_Caucasian_female_neutral_frontal
Rafd090_57_Caucasian_female_neutral_frontal
Rafd090_58_Caucasian_female_neutral_frontal
Rafd090_61_Caucasian_female_neutral_frontal

Male

008_y_m_n_a
016_y_m_n_a
031_y_m_n_a
037_y_m_n_a
049_y_m_n_a
081_y_m_n_a
089_y_m_n_a
099_y_m_n_a
105_y_m_n_a
114_y_m_n_a
127_y_m_n_a
147_y_m_n_a
160_y_m_n_a
21M_NE_C
24M_NE_C
am03nes
am04nes
am05nes
am06nes
am07nes
am08nes
am09nes
am10nes
am11nes
am14nes
am17nes
am18nes
am22nes
am25nes
am26nes
am28nes
bm12nes
bm16nes
M3neut_c_st
M7neut_c_st
Rafd090_03_Caucasian_male_neutral_frontal
Rafd090_05_Caucasian_male_neutral_frontal
Rafd090_07_Caucasian_male_neutral_frontal
Rafd090_09_Caucasian_male_neutral_frontal
Rafd090_10_Caucasian_male_neutral_frontal
Rafd090_20_Caucasian_male_neutral_frontal
Rafd090_23_Caucasian_male_neutral_frontal
Rafd090_25_Caucasian_male_neutral_frontal
Rafd090_36_Caucasian_male_neutral_frontal
Rafd090_49_Caucasian_male_neutral_frontal

Visual working memory

Female

02F_NE_C
054_y_f_n_a
063_y_f_n_a
085_y_f_n_a
098_y_f_n_a
101_y_f_n_a
125_y_f_n_a
150_y_f_n_a
152_y_f_n_a
162_y_f_n_a
173_y_f_n_a
182_y_f_n_a
af06nes
af07nes
af14nes
af15nes
af17nes
af19nes
af20nes
af22nes
af24nes
af28nes
af30nes
af35nes
F17neut_c_st
F23neut_c_st
Rafd090_12_Caucasian_female_neutral_frontal
Rafd090_16_Caucasian_female_neutral_frontal
Rafd090_18_Caucasian_female_neutral_frontal
Rafd090_19_Caucasian_female_neutral_frontal

Male

013_y_m_n_a
025_y_m_n_a
057_y_m_n_a
062_y_m_n_a
066_y_m_n_a
109_y_m_n_a
123_y_m_n_a
135_y_m_n_a
153_y_m_n_a
am01nes
am02nes
am13nes
am21nes
am27nes
am29nes
am31nes
am33nes
am34nes
am35nes
M16neut_c_st
Rafd090_15_Caucasian_male_neutral_frontal
Rafd090_21_Caucasian_male_neutral_frontal
Rafd090_24_Caucasian_male_neutral_frontal
Rafd090_28_Caucasian_male_neutral_frontal
Rafd090_30_Caucasian_male_neutral_frontal
Rafd090_33_Caucasian_male_neutral_frontal
Rafd090_38_Caucasian_male_neutral_frontal
Rafd090_46_Caucasian_male_neutral_frontal
Rafd090_47_Caucasian_male_neutral_frontal
Rafd090_71_Caucasian_male_neutral_frontal

Chapter IV

Study 3: Incidental learning of faces during threat: No evidence for increased autonomic arousal to “unrecognized” threat identities.

An adapted version of this chapter has been submitted to Scientific Reports and is currently under review as “Schellhaas, S., Schmahl, C., & Bublatzky, F. (under review). Incidental learning of faces during threat: No evidence for increased autonomic arousal to ‘unrecognized’ threat identities.”

4.1 Abstract

Remembering an unfamiliar person and the contextual conditions of that encounter is important for adaptive future behavior, especially in a potentially dangerous situation. Initiating defensive behavior in the presence of former dangerous circumstances can be crucial. Recent studies showed selective electrocortical processing of faces that were previously seen in a threat context compared to a safety context, however, this was not reflected in conscious recognition performance. Here, we investigated whether previously seen threat-faces, that could not be remembered, were capable to activate defensive psychophysiological response systems. During an encoding phase, 50 participants with low to moderate levels of anxiety (partially inpatients) viewed 40 face pictures with neutral expressions (6s each), without an explicit learning instruction (incidental learning task). Each half of the faces were presented with contextual background colors that signaled either threat-of-shock or safety. In the recognition phase, all old and additional new faces (total of 60) were presented intermixed without context information. Participants had to decide whether a face was new or had previously been presented in a threatening or a safe context. Results show moderate face recognition independent of context conditions. Startle reflex and skin conductance response (SCR) were more pronounced for threat compared to safety during encoding. For SCR, this differentiation

was enhanced with higher levels of depression and anxiety. There were no differential startle reflex or SCR effects during recognition. From a clinical perspective, these findings do not support the notion that perceptual biases and physiological arousal directly relate to threat-associated identity recognition deficits in healthy and clinical participants with anxiety and trauma-related disorders.

4.2 Introduction

Humans have difficulties in identifying unfamiliar persons, especially when the situational context changed (Davies & Milne, 1982; Burton & Jenkins, 2011; Young & Burton, 2017). This can become dangerous, e.g., when a person encountered in a hostile situation is met again but not recognized. During initial encounter, defensive response programs are activated when an evolutionary prepared aversive situation occurs (e.g., fear, freezing, fight/flight; Öhman & Mineka, 2001). In the second encounter, when the unfamiliar person is not identified, the aversive context of the first encounter is hardly recognized at all (Arnold et al., 2021). However, the question remains whether the unrecognized person is nevertheless associated with the threatening context even though neither is remembered. Thus, the present research question is whether we find defense activation (i.e., enhanced skin conductance response [SCR] and startle reflex) to non-recognized persons who were previously encountered in a threatening context.

Psychophysiological defense mechanisms have been suggested to operate relatively independently from conscious cognitive processing of stimulus associations (Öhman & Mineka, 2001). For instance, on the neural level we observed selective processing of previously seen faces as a function of whether faces were encoded in a threatening or safe context (Schellhaas et al., 2020). Interestingly, this differential old/new-ERP effect was not reflected in conscious recognition performance and resulted in poor face and context memory. Such difficulties in recognizing unfamiliar faces seen only briefly (e.g., for 1s or 6s) have also been observed in other studies in which a large number of ninety faces were presented in a blocked manner (Arnold et al., 2021). Participants did not know whether and, possibly more importantly, under what conditions they had seen a person before (i.e., during threat or safe conditions). Because neural processing is assumed to have direct access to defensive response programs irrespective of conscious face and context recognition (e.g., Carretié et al., 2009) the question emerged whether threat-selective perceptual processing results in priming of defensive psychophysiological response patterns to unrecognized threat relative to safety associated faces.

From a clinical perspective this is an important question as perceptual biases and physiological arousal may contribute to threat-associated identity recognition deficits in (socially) anxious and/or traumatized participants (Ehlers & Clark, 2000; Schellhaas et al., under review [a]). In this context, recent research suggested attentional threat biases and hypervigilance for highly anxious individuals (Shackman et al., 2016; Robinson et al., 2012). Moreover, the generalization of the defensive startle reflex activity is a central feature of pathological anxiety especially for posttraumatic stress disorders (Lissek et al., 2008; McTeague & Lang, 2012; Lang & McTeague, 2009). However, it remains unclear whether threat-related overgeneralization and attentional biases extend to the memory processes and remain constant once the acute threat has disappeared.

The present study examined whether participants show threat-enhanced activity of the autonomic and somatic nervous system (i.e., enhanced SCRs and threat-potentiated startle reflex) to people that were previously met in a threat context but cannot be remembered. To this end, an item/source (i.e., face/context) recognition task was combined with the threat-of-shock paradigm, in which participants were verbally informed about the possibility to receive electrical shocks when a particular colored background was present (e.g., blue signals threat) whereas another background color indicated safety (e.g., green signals safety). During an initial encoding phase (incidental learning), unfamiliar faces were paired with either a threat or safety context. It is assumed that defensive responses are increased for faces in a threat relative to a safe context (i.e., threat-potentiated startle and SCR; Grillon & Charney, 2011; Bublatzky et al., 2013). Moreover, startle potentiation should be more pronounced in more (socially) anxious participants while the difference between threat and safety decreases (generalization, Lissek et al., 2008; Lobo et al., 2015).

During the following recognition phase, an unexpected face/context recognition task was performed (Schellhaas et al., 2020; Arnold et al., 2021). Participants indicated whether they had seen a face with a threat or safety background, or whether it was new. Despite the notion of a memory enhancing effect of aversive apprehension (Dolcos et al. 2012; Ventura-Bort et al., 2016a), we expected rather poor recognition performance for incidentally learned faces (Schellhaas et al., 2020; Arnold et al., 2021; Burton & Jenkins, 2011). Of particular interest is the physiological responding to faces that had been seen previously within threat compared to safety background. According to the notion of a threat advantage (Öhman & Mineka, 2001), faces encoded during threat should elicit enhanced SCRs and startle reflex potentiation during recognition relative to safely encoded faces. Similar to encoding, more pronounced SCR and startle reflex are expected for (socially) anxious (Wangelin et al., 2012) and traumatized participants with more severe anxiety (Lobo et al., 2015). For these participants, less face differentiation (old-threat vs. old-safe vs. new) was expected, reflecting maladaptive evaluative processing of threatening face–context compounds.

4.3 Methods

Participants

Fifty participants (35 female, 12 male, 3 other) between the age of 18 to 51 years (M = 24.87 [SD = 7.46]) were recruited from the University of Mannheim, the SRH Heidelberg and the general population of Mannheim (Germany). For recruitment of a diverse sample, a trans-diagnostic and dimensional approach was followed. To increase the number of high-anxious participants with clinically relevant psychopathology, thirteen participants were recruited from the Department of Psychosomatic Medicine and Psychotherapy at the Central Institute of

Mental Health in Mannheim. Of all participants, 45.8% self-reported a current or past mental disorder. 31.3% of these participants self-reported current or past borderline personality disorder (BPD), 29.2% posttraumatic stress disorder (PTSD), 25% depression and 8.3% anxiety disorders. 14.3% self-reported another mental disorder (including attention deficit/hyperactivity disorder, (atypical) anorexia nervosa, obsessive-compulsive disorder, narcissistic personality disorder).

We assessed anxiety, depression and trauma measures (STAI-S = 38.88 [SD = 13.23], STAI-T = 49.68 [SD = 13.93], SPIN = 40.41 [SD = 15.68], BDI II = 41.03 [SD = 16.57], CTQ = 51.71 [SD = 22.72], using the German versions of the Social Phobia Inventory, SPIN, , Stangier & Steffens, 2001; State-Anxiety Inventory, STAI-T/S, Laux et al., 1981, Beck Depression Inventory II, BDI II, Hautzinger et al., 2006, Childhood Trauma Questionnaire, CTQ, Bernstein et al., 1998). Participants received monetary compensation for participation (24€). Due to technical issues, behavioral and physiological data from 9 participants (ID 1-9) were lost for the recognition phase and questionnaire measures are missing from 5 participants (ID 1-2, 4, 22 and 30). Six participants (ID 8 and 21 for encoding, and ID 21, 27, 44, 47 and 50 for recognition) were excluded from startle analysis due to poor data quality (i.e., excessive EMG artifacts) and two from SCR analyses (ID 8 and 30) due to non-responding. All participants provided informed written consent to the experimental protocol which was approved by the ethics committee of the Medical Faculty Mannheim, Heidelberg University (Germany) and complies with the APA ethical standards and the Declaration of Helsinki.

Sample size was determined based on previous studies using a similar threat-of-shock procedure (e.g., Bradley et al., 2018; Dunning et al., 2013) and estimation of sample size using G*Power analyses (Faul et al., 2007) assuming a medium effects size (f = .2), power (1-ß = .9), and correlation among repeated measures (r = .05) suggesting N = 48 to find reliable effects regarding item/context effects for the main startle analyses.

Materials and task

Sixty faces (30 female) with neutral facial expressions were selected from the Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998) and the Radboud Faces Database (RaFD; Langner et al., 2010). The pictures were processed with a photo editor (Adobe Photoshop CS2) and cropped to the same size (442×606 pixels) to reduce differences between the datasets. The stimulus set was randomly divided into three subsets of 20 face actors each (Set A, B, and C). During an encoding session, pictures of Set A and B were presented in a random order. Specifically, Set A was presented with a specific colored background frame (e.g., blue; RGB values: 0,255,0), whereas Set B was presented with a different colored frame (e.g., green; RGB values: 0,0,255; 1280×1024 pixels; see Figure 11). Assignment of subsets to background color was counterbalanced across participants. Face-context compounds were presented in a trial-to-trial pseudorandom order with the restriction of no more than two repetitions of the same background color. Pictures were presented for 6 s with a flexible ITI of 8-12 s. To provoke the defensive startle reflex, 10 auditory startle probes (104 dB, white noise) were administered per picture set, with a variable onset at 4.5-5.5s after picture onset (3 startle during ITI). To examine incidental learning and in order to keep recognition performance comparably low to our previous studies (Schellhaas et al., 2020; Arnold et al., 2021), the recognition task was not mentioned to the participants before the encoding session.

During the recognition session, all pictures (old Sets A and B, and new Set C) were presented intermixed without background colors and a combined item/source memory task was performed. Participants indicated by button press whether the face was previously seen (old) and if so in which context, or whether the face was new. They were instructed to guess the context if they remembered the face but not the background color, instead of choosing new. In total, 30 startle probes were equally distributed to the conditions old-threat, old-safe, and new faces; 3 additional probes were presented during ITI.

Stimuli were presented on a 22-inch computer screen placed approximately 0.8m in front of the participants using Presentation software (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Procedure

Participants completed the questionnaires on general and social anxiety as well as trauma history one to two days before their lab appointment, where the state version of the STAI was administered a second time. They were then seated in the experimental room and the electrodes for the physiological measures were attached. Following, participants rated the two background colors regarding valence and arousal using the Self-Assessment Manikin (SAM; Bradley & Lang, 1994, on a scale from 1-9) as well as perceived threat using a Likert scale (from *not at all* 0 to 10 *highly threatening*) which served as a baseline rating. Afterwards a habituation phase served to familiarize participants with the auditory startle probes (8 trials, not analyzed).

Next, a (fake) shock electrode was placed on the non-dominant forearm. The participants were then verbally instructed that they could receive up to three maximal unpleasant but not yet painful electric shocks while one background frame was present (e.g., blue serves as threat context) but not while the other was present (e.g., green serves as safety context). Verbal threat-of-shock instructions were used to investigate aversive expectation instead of direct experience (Olsson & Phelps, 2004). Color assignment to threat and safety was counterbalanced across participants. In the following encoding session, participants were instructed to attentively watch all presented face. After encoding, the shock electrode was detached, and participants rated the context colors regarding valence, arousal, and perceived threat as a follow-up measure. The recognition task followed immediately (Figure 11). Afterwards participants completed a short interview and were debriefed.

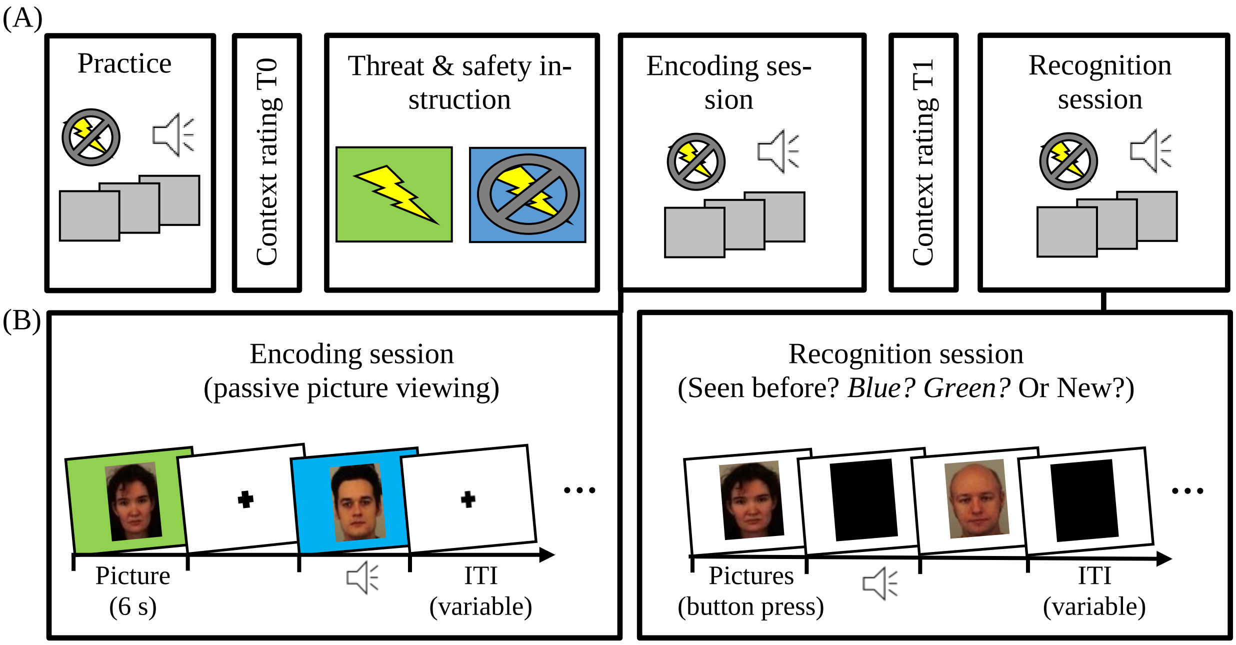


Figure 11. Schematic illustration of the experimental procedure. (A) Following initial practice trials, context colors were rated regarding valence, arousal and threat (T0) and then threat-of-shock and safety based on the colors were verbally instructed. Afterwards, an encoding session started, after which the context colors were rated a second time (T1). Finally, participants performed a recognition session. (B) During the encoding session 40 faces with neutral facial expressions were presented in front of either of the two alternating background colors (i.e., threat and safety) and the participants were instructed to view each face attentively. A startle sound (white noise) was presented for half of the trials. In the recognition session participants saw the 40 faces from encoding and 20 new ones and had to indicate by button press in front of which color they had been presented during encoding or whether they were new. Again, half of the trials were accompanied by a startle sound.

Data recording and reduction

Electrodermal activity (EDA), and electromyography (EMG) were recorded (1000Hz sampling rate) as indicators for autonomic nervous system activation using Ag-AgCl electrodes, AcqKnowledge software and Biopac amplifiers (BIOPAC Systems; Golet, CA). Additionally, heart rate (ECG) was assessed, but not analyzed. For skin conductance, electrodes were placed at the hypothenar eminence of the non-dominant hand. Skin conductance data was down sampled to 20Hz offline and noise was attenuated using Butterworth Zero Phase 2Hz low- and a 0.05 Hz high-pass filter. For the startle EMG, two electrodes were attached below the right eyelid, measuring the electromyogram of the orbicularis oculi muscle (Blumenthal et al., 2005). During recording frequencies below 28Hz and above 500Hz were filtered out with a bandpass filter (24 dB/octave roll-off) and a 50Hz Notch filter was applied to remove noise from the power line offline. This data was then rectified and smoothed with a moving average procedure (50 ms) in BrainVision Analyzer 2.0 (BrainProducts, Munich, Germany).

Startle responses were scored with a semi-automated procedure as maximum peak in the 21 – 150 ms time window following each startle probe. Peak amplitude was calculated relative to a mean baseline period (50 ms preceding startle response time window; cf. (Bublatzky et al., 2014). Skin conductance response (SCR) to startle probe and picture onsets were calculated as the maximum increase in skin conductance in the interval of 1 to 4.5 s (relative to a 2 s pre-stimulus period). A minimum threshold of 0.02 µS was used for zero-response detection, and range and distribution correction were applied.

Data analysis

The datasets generated during and/or analyzed during the current study are available in the OSF repository: https://osf.io/8afpm/?view_only=57d8e169e60a42e7b5d2ecfebff57170

Self-report data. Valence, arousal, and threat ratings were submitted to a repeated measure ANOVA with Time (T0 baseline vs. T1 follow-up) and Context (threat vs. safety) as within-subjects factors. Helmert contrasts were used to follow-up on main effects at a significance level of $p < .05$. Greenhouse-Geisser corrections were applied where necessary and as a measure of effect size, partial eta squared (η_p^2) is reported.

Behavioral data. Regarding recognition performance, hierarchical multinomial processing tree modeling (MPT; Heck et al., 2018) was used providing parameter estimates for item- and source-memory as well as guessing parameters. Moreover, this approach enables examining co-variations between memory parameters and inter-individual questionnaire measures like anxiety and depression scores. The two-high-threshold source monitoring model (2HTSM, Bayen et al., 1996) for two sources was used.

Three discrete answer categories (threat context "T", safe context "S", and new "N") are assumed to be derived from certain latent cognitive processes, modelled as probabilities in three decision trees (see Figure 7). Specifically, item recognition (D_T, D_S & D_N), source recognition (d_t, d_s), item guessing (b), and source guessing (g). The first decision tree represents the case of an item presented with a threatening context. With probability D_T the item is recognized as being previously seen (i.e., old) and with the probability d_t the source (i.e., threatening context) is also recognized. If the item is not recognized as old ($1-D_T$), participants correctly guess with probability b that the item was old. In that case, the source must be guessed as well, and with probability g the correct source is chosen (i.e., threat). The second decision tree represents the case of an item previously presented with a safe context and the probabilities derive accordingly. The third decision tree represents a previously unseen (i.e., new) item. With probability D_N the item is recognized correctly as being new, with probability $1-D_N$ it is falsely classified as old and the parameters arise accordingly to the first two decision trees. In this current form, the model is not identifiable. Therefore, the parameters for recognizing old safe faces and new faces were restricted to be equal ($D_S = D_N$), while D_T remained unrestricted. These restrictions are typical of the 2HTSM (e.g., Bayen et al., 1996; Bell et al., 2017). Additionally, guessing parameters were set equal (b = g), under the assumption that guessing tendencies would be the same regardless of item recognition (Arnold et al., 2013; Bell et al., 2017). Analyses were conducted using the TreeBUGS package (Heck et al., 2018) for R (R Core Team, 2016), questionnaire data were entered as covariates into the model.

Additionally, hit rates (HR), false alarm rates (FAR), item recognition (HR – FAR) and the conditional source identification measure (CSIM, number of correctly identified source attributions divided by the total number of targets, Arnold et al., 2021) were calculated and submitted to an ANOVA with the factor Context (threat vs. safety).

Physiological data. Startle EMG and SCR were submitted to repeated measures ANOVAs with the factor Context (threat vs. safety), separately for the encoding and recognition phase. For SCR the factor of Block (first half vs. second half) was included to account for habituation effects.

In order to explore the probability of null effects of our central hypotheses, we conducted Bayesian analyses (Kass & Raftery, 1995), using the R based software package JASP 0.16.1 (JASP Team, 2022). Thereby we focused on the central effect of Context on behavioral performance, startle potentiation, and SCR. Additionally, for SCR the factor Block and the interaction Context × Block were included. Using Monte-Carlo sampling 10,000 iterations and default prior scaling factors (for fixed effects = 0.5, random effects = 1, r covariates = 0.354, Bayes factors (BF) were estimated (Rouder et al., 2012). BF inclusion scores (BF_{incl}) are reported that inform about how much the inclusion of the factor (e.g., Context) is supported by the data, compared to the null-model. A value below 1 (above 1) indicated that the data is more likely for the null-hypotheses than for the alternative hypotheses (and vice versa).

Greenhouse-Geisser corrections were applied where necessary, and the partial eta square (η_p^2) is reported as a measure of effect size. To control for type 1 error, Bonferroni correction was applied for post hoc t tests.

4.4 Results

Self-report data

The rating data revealed an overall successful subjective induction of threat anticipation (see Figure 12). For the valence ratings, neither the two contexts nor the two time points T0 (baseline) and T1 (follow-up) were rated differently, Time $F(1,47) = .38$, $p = .54$, $\eta_p^2 = .01$, Context $F(1,47) = 1.35$, $p = .25$, $\eta_p^2 = .03$. A significant Time × Context interaction showed, however, that the safety context was perceived as more pleasant than the threat context after threat induction but not at baseline, Time × Context $F(1,47) = 29.39$, $p < .001$, $\eta_p^2 = .39$. The overall level of arousal was higher during T1 than during T0, Time *F*(1,47) = 10.03, $p < 0.01$, $\eta_p^2 = .18$. While in general the contexts did not differ, Context *F*(1,47) = 2.32, $p = .13$, $\eta_p^2 = .05$, post-hoc tests revealed that, after the threat induction, the threat relative to safety context was rated more arousing, Time × Context *F*(1,47) = 27.18, $p < .001$, $\eta_p^2 = .37$. Threat levels increased over time, Time *F*(1,47) = 20.98, $p < .001$, $\eta_p^2 = .31$ and in intensity, Context *F*(1,48) = 4.10, $p = .05$, $\eta_p^2 = .08$, with post-hoc tests showing that the threat context was perceived as more threatening than the safe context only during T0 and not during T1, Time × Context *F*(1,47) = 26.22, $p < .001$, $\eta_p^2 = .36$. There were no covariation effects with questionnaire scores on any of these measures.

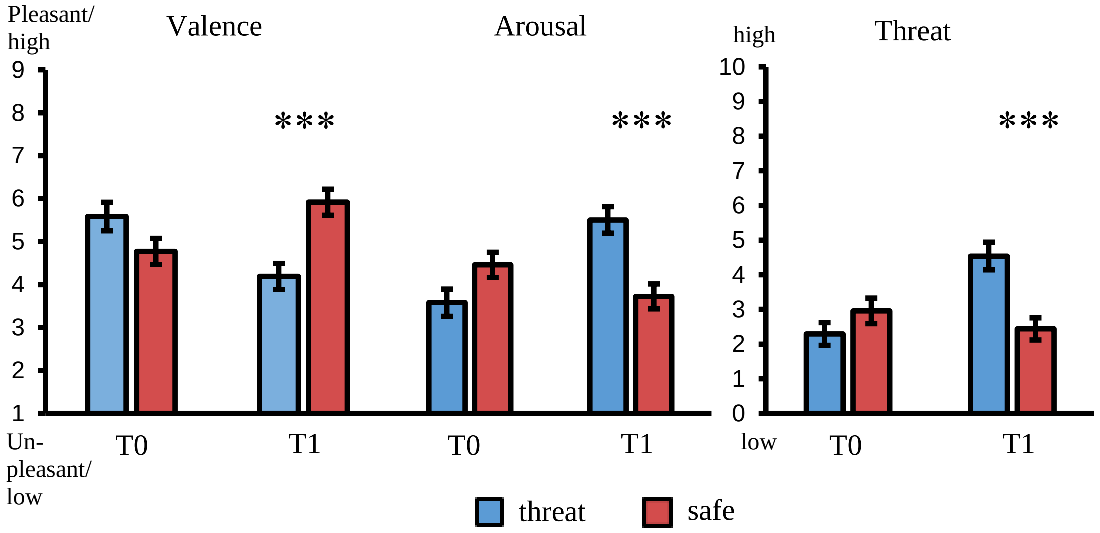


Figure 12. Mean valence, arousal (scale of 1 to 9) and threat (scale of 0 to 10) ratings of the instructed threat and safety contexts, indicated by the colors blue and green (random assignment) before the threat instruction (T0) and after the encoding session (T1) (M and SEM, *** p < .001).

Behavioral data

Table 6 contains the parameter estimates of the hierarchical MPT. Using Bayesian modelling, a priori distribution of the parameters is used before analyses and compared to a posterior distribution which is based on the data using Baye's theorem. The Bayesian confidence intervals

(BCIs) explains within which range the true parameter lies in the posterior distribution, given data and prior distribution. Overlapping BCIs indicate non-significant differences between parameter estimates. While participants were moderately able to recognize items as old or new, it did not matter whether they had seen them in a threatening or a safe context during encoding. This is indicated by the overlapping BCIs of the D_T and D_S parameters.

Moreover, source recognition was poor ($d_T = .04$ and $d_S = .01$ with large Bayesian confidence intervals for the threatening source [.00 - .39]) and did not differ between a threatening or safe source from encoding, evident in overlapping BCIs of the d_t and d_s parameters. The guessing parameters reflected the probabilities of the events for source guessing ($g = .48$ with half of the faces from a threatening and half of the faces from a safe context). The probability of a face having been presented before during encoding compared to new was 2:1. There was a slight conservative guessing tendency of classifying an old face as a face being new ($b = .58$ with the actual probability of a face being old of .67). No significant associations emerged between parameter estimates and questionnaire scores, which were entered as covariates into the model. Model fit (Klauer’s [2010] test statistics $T1$ $p = .25$) was good.

Using conventional recognition parameters, mean hit rates and false alarm rates for the threatening context were HR $M_{threat} = .76$ ($SD = .15$), FAR $M_{threat} = .20$ ($SD = .17$) and for the safe context HR $M_{safe} = .73$ ($SD = .16$) and $M_{safe} = .20$ ($SD = .15$). Item recognition (HR – FAR) did not differ based on Context, $F < 1$, $p = .40$, $\eta_p^2 = .02$, $BF_{incl} = 0.305$. Likewise, source identification did not differ significantly as a function of context conditions, Context $F < 1$, $p = .50$, $\eta_p^2 = .01$, $BF_{incl} = 0.290$ ($M_{threat} = .21$ [$SD = .07$], $M_{safe} = .22$ [$SD = .05$]). This means that it is 3.28 and 3.45 times more likely that there is no difference between the two context conditions for both behavioral measures. None of the questionnaire measures had an impact on memory performance, $Fs < 1$, $ps > .73$.

Table 6. *Mean parameter estimates of the latent-trait MPT model for the recognition performance of the item/source memory task.*

Parameter	M [95 % BCI]	SD
g	0.48 [0.45 – 0.52]	0.02
b	0.58 [0.49 – 0.67]	0.05
d_T	0.04 [0.00 – 0.39]	0.11
D_T	0.28 [0.05 – 0.48]	0.11
d_S	0.01 [0.00 – 0.04]	0.02
$D_S = D_N$	0.41 [0.32 – 0.49]	0.04

Note. For the group-level estimates, posterior means (and SDs) are shown. BCI = Bayesian confidence interval. D_T, D_S and D_N =face recognition parameters, d_T, d_S = context memory parameters and b, g = guessing probabilities.

Startle reflex

As expected for the encoding phase, the startle reflex was potentiated for faces presented with the threatening compared to the safe background, Context $F(1,45) = 37.20$, $p < .001$, $\eta_p^2 = .45$ (see Figure 13A). A marginally significant Context × Depression interaction indicated that the threat-potentiated startle was more pronounced for more depressed participants, $F(1, 41) = 3.76$, $p = .06$, $\eta_p^2 = .08$. No significant interaction emerged for other questionnaire measures.

Regarding the recognition phase, no difference in startle potentiation was found, Context $F(2,70) < 1$, $p = .64$, $\eta_p^2 = .01$, $BF_{incl} = 0.133$, making the null hypotheses 7.52 times more likely than the alternative hypothesis. Individual differences based on the questionnaire scores did not modulate the findings.

Skin conductance responses to startle probes

Skin conductance responses to startle probes were more pronounced during the first relative to the second half of the encoding session, Block $F(1,43) = 31,43)$, $p < .001$, $\eta_p^2 = .42$. No main effect of Context emerged, $F < 1$, p = .53, $\eta_p^2 = .01$, however a significant interaction Context

× Block, $F(1,43) = 6.56$, $p < .05$, $\eta_p^2 = .13$. Post-hoc tests revealed that the threatening context elicited a more pronounced SCR compared to the safe context only during the second half of the encoding session, $F(1,43) = 6.83$, $p < .05$, $\eta_p^2 = .14$. Moreover, (marginally) significant interactions emerged for Context with depression, trait-anxiety, and social anxiety, *Fs*(1, 40) = 4.01, 4.93, and 6.02, *ps* = .05, and < .05, $\eta_p^2 = .09$, .11, and .13. These interactions indicate more pronounced differentiation between threat and safety context with higher depression and anxiety scores.

For the recognition phase, startle locked SCRs were not modulated by Context, $F < 1$, $p = .45$, $\eta_p^2 = .03$, $BF_{incl} = 0.074$, or Block $F(1,28) = 3.21$, $p = .08$, $\eta_p^2 = .10$, $BF_{incl} = 0.995$. There was also no significant Context × Block interaction, $F < 1$, $p = .45$, $\eta_p^2 = .03$, $BF_{incl} = 0.04$, making the respective null hypotheses 13.51, 1.01, and 25 times more likely than the alternative hypotheses.

Skin conductance responses to picture onset

During the encoding session, SCRs locked to picture onset varied as a function of Context, $F(1,43) = 29.93$, $p < .001$, $\eta_p^2 = .41$, and Block, $F(1,43) = 60.55$, $p < .001$, $\eta_p^2 = .59$, indicating more pronounced SCRs during threat compared to safety, and during the first compared to the second half of the encoding session (see Figure 13B). There was no significant interaction Context × Block, $F(1,43) = 2.10$, $p = .16$, $\eta_p^2 = .05$, nor any interactions with the questionnaire measures, *Fs* < 1, *ps* > .43.

During recognition, SCRs elicited did also not vary between threat and safety, Context, $F(2,72) = 1.05$, $p = .35$, $\eta_p^2 = .03$, $BF_{incl} = 0.067$, nor between first and second phase of the recognition session, Block $F < 1$, $p = .83$, $\eta_p^2 = .00$, $BF_{incl} = 0.106$. The Context × Block interaction was not significant, $F(2,72) = 1.53$, $p = .22$, $\eta_p^2 = .04$, $BF_{incl} = 0.011$. Thus, the respective null hypotheses were 14.93, 9.43, and 90.91 times more likely than the alternative hypothesis. Taken together, in the recognition phase, there was no indication that autonomic arousal was enhanced for faces that had been previously presented with a threatening compared to a safety context. There were no significant influences of the questionnaire measures, *Fs* < 1, *ps* > .51.

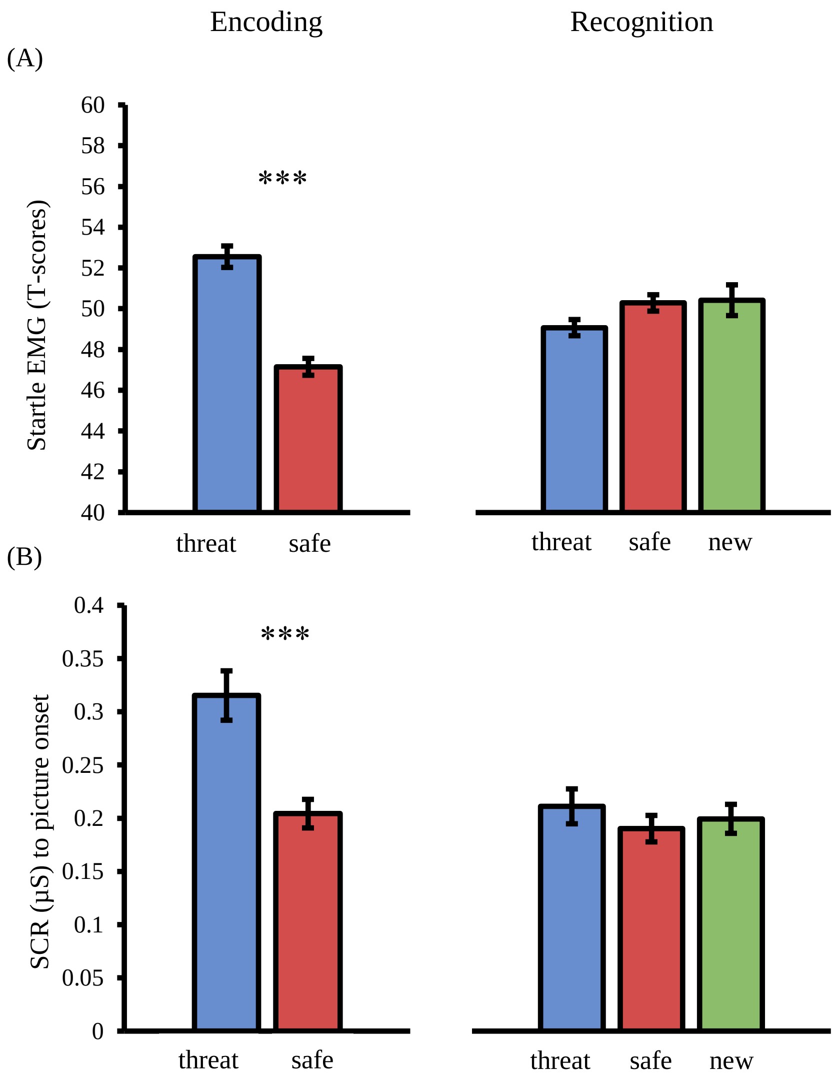


Figure 13. (A) Mean startle reflex EMG and (B) skin conductance response to picture onset for the encoding and recognition session, as a function of contextual threat and safety, or new faces (M and SEM, *** p < .001).

4.5 Discussion

Previous research has shown that encoding unfamiliar faces within an aversive anticipatory context (i.e., threat of shock) leads to later threat-selective neural processing (indicated by event-related brain activity), although recognition performance was low (Schellhaas et al., 2020; Schellhaas et al., under review [a]). In the present study, we examined the hypotheses whether autonomic arousal can be found for such threat-associated but non-recognized unfamiliar face pictures. Psychophysiological measures of the startle reflex and SCR did not support this hypothesis. Aversive apprehensions were successfully induced via verbal threat instruction in the encoding session and resulted in a more unpleasant, arousing, and threatening rating of the threat compared to the safety context. Moreover, similar to previous research, a threatening context activated the autonomic and somatic nervous system during the encoding session, as evidenced by threat-enhanced skin conductance responses and potentiated startle reflexes compared to the safety context (e.g., Bublatzky et al., 2022). As expected for an implicit learning task, memory for unfamiliar faces was rather poor and did not vary as a function of contextual threat or safety Schellhaas et al., 2020; Arnold et al., 2021). In fact, participants were barely able to recognize whether they had seen a face before, and thus were unable to identify the contextual situation (i.e., threat or safety source) of a previous encounter. Similarly, autonomic and somatic activation did not vary during the recognition session, neither for faces from a threat or safety source, nor for newly presented faces. Moreover, different measures of psychopathology (i.e., depression, social and state/trait anxiety, early maltreatment) did not affect psychophysiological responding to faces during the recognition session. These findings are further supported by Bayesian analyses, showing that the null hypotheses in the recognition phase (i.e., no difference between faces from a threat or safety context or new faces) were much more likely than the alternative hypotheses.

A dangerous situation triggers the activation of defensive psychophysiological systems in order to avoid or, if a flight reaction is not possible, at least minimize harm (Gilbert, 2001). Such response programs to danger cues include the activation of the autonomic and somatic nervous system and can persist even when the dangerous situation has evidently passed. This is particularly true when the threat was learned through social means such as verbal instructions

(Mertens & De Houwer, 2016; Mertens et al., 2018). The pure anticipation of threatening events (e.g., unpleasant electric shocks; Grillon et al., 1991) is sufficient to provoke a relatively persistent state of aversive apprehension even when the threatening situation never occurred (e.g., across repeated test days without shock reinforcement; Bublatzky et al., 2014, 2022).

However, once the threat has been averted and its signals have disappeared, a conscious memory of people and situations associated with the threat seems necessary to be better prepared for future encounters (Lang et al., 2000). While we did not find support for this hypothesis in the present study design, several aspects and alternative hypotheses need to be considered. First, the removal of the shock electrode after the encoding phase together with the instruction that no electrical shock can follow after this point may have terminated the real threat (instructed extinction; Luck & Lipp, 2016). Second, the complete absence of threat cues during recognition probably triggered a general downregulation of defensive behaviors (Rowles et al., 2012). However, compared to the startle reflex activity during a safe encoding context, no reduction or downregulation was observed in the recognition session for formerly threatening, safe or new faces. Third, aversive apprehensions have been suggested to positively improve cognitive performance (Dolcos et al., 2012). For example, arousal improves accuracy in a sustained attention to a response task (Aylward et al., 2017) or in recognizing neutral objects from emotional backgrounds (Ventura-Bort et al. 2016a). In another study, memory of context (arousing vs. non-arousing) was found to be impaired by the presence of negative information, whereas threat of shock enhanced item memory (Bisby & Burgess, 2014). Therefore, negative affect appears to impair memory for associations, whereas storage of negative perceptual representation is spared or even enhanced.

Fourth, in the present study, the facial stimuli were prominently placed in the foreground, while contextual threat and safety colors served as a background. Although participants were instructed to attend to every picture, the colored background indicated threat of electrical shocks or safety and was relevant for adaptive behavioral response priming (e.g., avoidance, heightened attention). In contrast, faces were a prerequisite for source identification during the recognition session (Bell et al., 2017). Together with a possible item-source dissociation during encoding, this might have led to an overall decrease in psychophysiological responding to a potentially dangerous situation (Schellhaas et al., 2020).

Although poor memory performance was intended in this study (to investigate the possible activation of defensive systems without conscious recollection), future studies should strengthen the item-source association during encoding to promote familiarity-based retrieval (Diana et al., 2011; Mather & Sutherland, 2011). Moreover, future studies are needed to investigate the effect of threat on recognized faces by increasing memory performance, for example by lowering the number of faces to be remembered. Here, explicit learning instructions and increasing the number of presentations per face could be pertinent. This procedure could be used to clarify whether defensive reactions are triggered even when there is no conscious recognition of previous threatening encounters.

Although the recognition of unfamiliar faces was expectedly poor (Burton & Jenkins, 2011), and the association of the faces with the threatening (and safe) context(s) not explicitly emphasized and strengthened, this study aligns with other findings that show the impact of contextual surroundings on the perception and processing of otherwise neutral faces (Wieser et al., 2014; Klein et al., 2015; Ventura-Bort et al., 2016a; Schellhaas et al., 2020). Extending these findings, we found enhanced SCRs to the onset of faces (and startle probes) within a threat compared to safety context, indicating enhanced arousal and alertness to these faces–context compounds. This effect diminished, when eliminating the threat-related part of the face-context associations during recognition, where SCRs were similar for all faces (i.e., threat, safe, and new), especially when being unable to remember the presented stimuli. This underscores the importance of the interplay of bottom-up (automatic) and top-down (expectation-based) factors in threat perception (Sussman et al., 2016; Schindler & Bublatzky, 2020; Bublatzky et al., under review). It also supports the claim that neutral facial expressions do not convey sufficient safety information and are disregarded in a dangerous context (Bublatzky et al., 2022). In addition, similar contexts during encoding and recognition have long been found to boost memory performance (Tulving & Thomson, 1973; Murnane et al., 1999). As the context conditions fundamentally changed from encoding to recognition, a cued recognition task with presenting the faces together with all contexts from encoding could have led to a renewed autonomic nervous system activation.

From a clinical perspective, the present study adds insight to several important points. While previous studies have shown that a threatening environment can change perception, attention, and memory processes (e.g., Schellhaas et al., 2020, under review [a]), here we found no indication for our key prediction that threat-encoded faces will lead to enhanced defense activation during recognition. However, psychophysiological responding to threat/safety situations varied with interindividual differences in depression and trait-anxiety measures during encoding of face–context compounds. Specifically, threat-enhanced skin conductance response increased with anxiety and depression scores. This differentiation did not persist into the recognition phase of the experiment, possibly due to response habituation. Alternatively, these findings are in line with previous research showing that anxiety, at least at low-moderate levels, has no significant impact on implicit memory and recognition and does not favor a memory bias towards threatening information (Mitte, 2008). Participants with moderate-high levels of anxiety and depression should be included in future studies to gain insight to the effect of former contextual arousal on face recognition.

In summary, a first encounter with an unfamiliar person that was combined with contextual threat leads to activation of defensive response systems (threat-potentiated startle reflex and SCRs during encoding). However, this effect disappeared during the second encounter with this person without contextual information (recognition session), indicating that the presence of contextual threat signals is necessary to trigger defensive responding, especially when memory of the threat–face association is poor.

Chapter V

General Thesis Discussion

This doctoral thesis is comprised of three empirical studies that set out to deepen the understanding of threat-related memory processes. Recognizing persons from formerly threatening encounters is a desirable skill and can be crucial for preventing bodily harm when physical integrity is at risk. However, environments are changing dynamically and can be ambiguous. The aim of this thesis was to identify the conditions that contribute to well-functioning processing of and behavior towards arousal in the context of (successful) memory performance. Additionally, the influence of potential maladaptive individual prerequisites such as anxiety and chronic stress exposure through childhood trauma was investigated.

To this end, in Study 1 thirty healthy participants performed an item/source memory task in alternating threatening and safe context conditions (i.e., source) during encoding of neutral faces (i.e., items). A threat advantage regarding perceptual processing as well as recognition performance for both faces and contexts was expected. It was hypothesized that arousal would increase the strength of mental representation and lead to an enhanced recognition of formerly unfamiliar faces. Associations to social anxiety were explored.

The results of Study 1 showed a dissociation between perceptual processing and conscious face recognition. Contrary to predictions, the ability to recognize previously presented unfamiliar faces was very poor, therefore no inference on source identification or the impact of threat and safety on face recognition could be drawn. However, neural processing differed during this recognition phase based on the differential encoding contexts of the faces. Thus, although the participants could not remember whether they had seen a face before, ERPs were modulated by arousing contexts, showing enhanced early and late negativity to threat in healthy participants. To gain more insight into the behavioral effect of arousal on memory, it was highly necessary to improve recognition performance.

In Study 2, a community sample of sixty-five participants with a history of ACE- and anxiety-related psychopathology, which was assessed on a dimensional level, performed the same memory task as in Study 1. Additionally, a visual working memory tasks, equally performed under threat and safety, was included to compare the effects of arousal on different levels of the information-processing stream and as a gateway to higher order cognitive functioning such as episodic memory (item/source memory). A maladaptive processing and disproportionate focus on threat was assumed with a threat bias in perceptual processing and impairments in memory functioning, especially for highly anxious and severely traumatized individuals.

Explicit learning instruction and a prolonged presentation time of the faces as well as grouping the faces in threat and safety blocks rather than varying the context conditions randomly and frequently, led to an overall improvement of face memory compared to Study 1. Faces that were previously associated with safety were recognized more easily than faces from a threatening context. This effect was stronger with increasing levels of anxiety and traumatization, indicating the importance of safety for undisrupted long-term memory processing. This could also hint to a possible attention-driven bias towards threat that leads to impaired memory encoding when arousal is task-irrelevant.

Substantial learning and memory deficits can accompany trauma experiences which was neither found in the acquisition of threat during social learning nor in the memory tasks in Study 2, where participants performed even better under the improved task conditions compared to Study 1. However, the corresponding contextual features (i.e., threat and safety) could not be reconstructed and the faces could not be attributed to the most probable source (Johnson et al., 1993), replicating the results of Study 1. This indicates that the task of memorizing and recognizing unfamiliar faces is already too challenging to be able to infer an association of the contexts (Johnston, & Edmonds, 2009).

Similar to the healthy participants and replicating findings of Study 1, there was an apparent dissociation in the lack of conscious recognition of the contextual settings and differential neural processing of threat and safety. This electrocortical differentiation was, however, reversed for threat and safety, indicating a potential marker for a threat bias after trauma exposure.

Study 3 was designed to test whether the absence of conscious perception of threat in formerly threatening settings would nevertheless lead to the activation of defensive systems. Since Studies 1 and 2 revealed a lack of an explicit memory trace for the context conditions, but differential neural processing of threat and safety, it was hypothesized, that the activation of the somatic and autonomic nervous system does not require conscious threat perception. Therefore, fifty-five participants completed the item/source memory task while psychophysiological responding (startle EMG, EDA) was measured.

The results of Study 3 revealed that conscious memory of individuals associated with past dangerous situations appears to be necessary to be better prepared for future, potentially equally threatening encounters. The participants did not show automatic arousal to faces that were met in a threat context but could not be remembered. The dissociation of perceptual processing and overt behavior, revealed in Study 1 and 2, did not translate into an activation of the somatic and autonomic nervous system when arousal was no longer present. Face recognition performance was moderate and independent of context. These findings do not support the notion that perceptual biases and physiological arousal directly relate to threat associated identity recognition deficits.

In the following section, I will integrate the findings into the research on the effect of socially transmitted arousal on perception and memory for face identity. Modulating effects of (clinical) anxiety and childhood maltreatment experiences will be discussed.

5.1 Integration of Study Findings and Research Implications

Arousing contexts through social learning. Aversive anticipation modulates perceptual processing and behavioral responding such as memory performance, decision-making, reward-learning (e.g., Baas, et al., 2002; Bublatzky et al., 2010; Bublatzky & Schupp, 2012; Bublatzky et al., 2017, Bublatzky, et al., 2020; Paret & Bublatzky, 2020; Weymar et al. 2013; Ventura-Bort et al., 2016b; Golkar & Olsson, 2017; Bublatzky, Schellhaas, & Paret, under review). Not only the effectivity but also the stability of socially transmitted aversive apprehensions has been consistently shown to modulate a variety of psychophysiological response systems as well as neural processing. Threat-potentiated activation of the somatic and autonomic nervous system has been demonstrated for instructed or observed threat relative to safety cues (e.g., enhanced skin conductance responses, heart rate deceleration, and potentiated startle responses; Bradley, et al., 2005; Bublatzky, et al., 2014, 2013, 2018, 2019; Costa et al., 2015; Grillon et al., 1991; Grillon & Davis, 1995; Funayama et al., 2001; Olsson & Phelps, 2004). Furthermore, threat-associated stimuli modulate and sensitize early sensory processing, evident in enhanced amplitudes of threat-cues in early ERPs mainly over visual processing areas and frontal distributions (P1, P2; Baas et al., 2002) as well as late threat-potentiated ERPs over parieto-occipital brain regions (LPP; Bublatzky & Schupp, 2012; Baas et al., 2002; Böcker et al., 2004). Functional magnetic resonance imaging (fMRI) shows that verbally transmitted or observed threat compared to safety leads to the activation of the amygdala with an involvement of the insula and ACC and a cortically distributed network with the hippocampus at its center, indicating similar mechanisms to those involved in direct fear conditioning (Olsson & Phelps, 2007; Olsson et al., 2007; Phelps et al., 2001).

The threat-enhanced perceptual and evaluative processing (Wieser & Brosch, 2012; Klein et al., 2015) through social learning pathways was replicated in all three studies of this thesis, indicating effective social threat learning. This was evident in all measures of psychophysiological responding (enhanced startle EMP and SCR to threat cues; Studie 3), electrocortical processing (early parieto-occipital negativity as a sign of early attentional tagging and late fronto-central negativity as an indicator of motivationally driven processing of threat- face compounds; Studies 1 & 2) as well as ratings which were more arousing, threatening and less pleasant for the threat-cue (Studies 1-3). While most of the studies familiarized the participants with the aversive stimulation by using a shock-work-up before the social learning procedure (Bublatzky et al., 2012; Bradley et al., 2005; Lindström et al., 2018), the adjustment to an individual level of uncomfortableness to the electric shock was omitted in all three studies. The results showed that the absence of prior direct experience to the aversive events equally established aversive anticipation, indicating the strength and persuasiveness of social interactions. It is therefore not necessary to face the dangerous stimuli, establishing social threat learning as a powerful and useful tool to study and induce state anxiety in a "safe" way in a laboratory setting for a wide variety of participant groups otherwise not suited for the use of actual (high intensity) electrical stimulation (e.g., physically impaired participants).

Further expanding on these findings, the social learning pathways worked equally well in establishing aversive apprehension in healthy participants in Studies 1 and 3 and in individuals with a history of childhood maltreatment in Study 2, being one of the few studies examining this effect in psychopathology (Debiec & Olsson, 2017; Robinson et al., 2013; Blair et al., 2016). There is practical and empirical evidence that vicarious fear learning can play an essential role in the development and maintenance of psychopathology, especially in anxiety- and trauma-related disorders and the social transmission of fear, for instance between a parent and child (DSM-5; American Psychiatric Association, 2013; Debiec & Olsson, 2017; Askew & Field, 2007, 2008; De Rosnay et al., 2006). Further, both the observation of fear in a model and the instruction of to-be-expected threat by the experimenter equally resulted in threat-enhanced neural processing. In general, similar mechanisms and outcomes are assumed for the different learning pathways (Olsson & Phelps, 2007; Olsson et al., 2007). However, while an US in association with a CS in a conditioning paradigm is assumed to have an universal impact on human beings in the sense of an evolutionary prepared response, various factors impact the learning quality in social fear learning, such as factors of the model and the instructor. For instance, group affiliation (through age, gender, race, etc.), perceived persuasiveness of the expression and skills of the model as well as empathy towards the model and the instructor influence how well and elaborated the information is learnt (Selbing et al., 2014; Golkar et al., 2015; Olsson et al., 2016). The relevance for further empirical investigation of (differential) effects of the social learning pathways in clinical populations is of great importance to identify possible risk factors and predictive factors for the development of maladaptive fear.

Impact of arousal on face perception, and short- and long-term memory for faces. The literature regarding the effect of emotional arousal on memory and cognitive functioning, including those based on verbal or observational fear learning, is extensive and often contains contradictory results (Robinson et al., 2013; Dolcos et al., 2020; Weymar et al., 2013). The release of stress hormones (Roozendaal et al., 2009; McGrath, 1976) and the resulting activation of stress-related brain networks centering around the amygdala are linked to enhanced connectivity to regions involved in learning and memory (hippocampus, prefrontal cortex; Hamann, 2001; Erk et al., 2003; Keightley et al., 2010). This is assumed to play a crucial role in emotionally enhanced memory over non-emotional stimuli (Kensinger, 2009). However, this effect seems to depend on various factors such as the time-point when arousal enters into the information-processing stream of memory (Kensinger et al., 2006; Baddeley et al., 2009) or individual differences such as anxiety or exposure to early and chronic stress experiences. Mixed findings regarding the emotion-enhancing memory effect have therefore been reported for different tasks and populations, with scarce evidence for neutral events incorporated in arousing scenes.

There is evidence on an impairing effect of stress and arousal (e.g., state-anxiety) on visual working memory (i.e., short-term memory) and its capacity (Robinson et al., 2013; Bergmann et al., 2012; Vytal et al., 2012; Lavric et al., 2003; Shackman et al., 2006; Lindström & Bohlin, 2012) with a contrasting beneficial effect on long-term and episodic memory (Robinson et al., 2013; Hamann, 2001; Dolan, 2002; Kensinger & Corkin, 2003). However, memory performance in the item/source (i.e., long-term) memory task in all three studies was poor. While face recognition in Study 3 was somewhat improved compared to Study 1, where participants were not able to recognize the faces as formerly met or new at all, no modulating effect of the arousing context to face memory was revealed. Furthermore, no context associations to previously seen faces could be retrieved in any of the three studies, concluding that there is no memory-enhancing effect of an arousing source neither on item nor on source memory itself, demonstrating a dissociation between arousing item and arousing source memory (Symeonidou & Kuhlmann, 2022; Bisby & Burgess, 2014; Bisby et al., 2018). One of the main driving factors for the lack of influence of arousal on recognition processes in these studies could be the inability to recognize unfamiliar faces in general (e.g., Arnold et al., 2021). Person identification for familiar and known faces is a well-developed skill in humans, whereas recognizing unfamiliar individuals is very challenging (Haxby et al., 2000; Johnston, & Edmonds, 2009; Bruce et al, 2001). The short presentation time of the faces in Study 1 and the constantly changing context conditions (i.e., trial-wise presentation of arousing and non-arousing contexts) made the task even more difficult. Additionally, in Study 1 and 3 participants were not prepared for a recognition test and viewed the face pictures during encoding passively together with the (non)arousing contexts. Insufficient integration of the faces into the context, that was more salient and carried higher relevance to the participants in the form of danger of electric shocks, could have led to an overall dismissal of the faces.

In Study 2, the conditions of person perception were improved by prolonging their viewing times and including explicit learning instructions as well as introducing arousal as a more permanent factor (i.e., block-wise presentation of the stimuli). Thereby, person identification was improved and slightly better for faces that were associated with a safe background. Source assignment to the faces, however, was still not possible. Compared to long-term-memory, visual working memory performance was entirely unaffected by arousing contextual conditions. Reduction of cognitive resources by increasing cognitive load led to a performance drop in the change detection task which was not modulated by the alternating threat and safe contexts. This could indicate a disregard of non-task-relevant information, even in the form of potentially harmful events in highly engaging tasks or a trade-off in efficiency and accuracy under threat. The latter was not supported by the (exploratory) results of reaction time in the working memory task, which was similar for threat and safety under high load. However, reaction time was not manipulated in the study and was explicitly stated by the task instructions to not play a role for the evaluation of task performance. For a more conclusive interpretation, time restraints on the working memory task seem necessary in future studies.

Following an Arousal Biased Competition approach (ABC; Mather & Sutherland, 2011), arousal was suggested to bind attentional resources by increasing the salience of a cue/stimulus. This can also lead to an enhancement of corresponding features such as color or location of the arousing item and in consequence to better memory both for the arousing stimulus itself but also its surroundings (Mather & Sutherland, 2011; Kensinger et al., 2006; Doerksen & Shimamura, 2001). However, if an only distinctly associated (e.g., temporarily or spatially close) stimulus is less salient, it must compete for remaining attentional resources, leading to reduced processing and decrease in memory (Mather & Sutherland, 2011; Bisby & Burgess, 2017). Therefore, the key factor for gaining a surplus of attention is the priority with which an information is processed (e.g., Ventura-Bort et al., 2016b). Attentional Control Theory (ACT; Eysenck et al., 2007) adds the notion that arousal reduces control over attentional processes, as salience of the arousing stimulus claims attention over top-down influences. If arousal is task-irrelevant, compensatory strategies can, however, counterbalance detrimental effects on performance and improve efficiency.

The behavioral findings of all three studies can be interpreted within predictions from ABC and ACT. Previous research reported an increase in memory for non-social stimuli, such as words or objects, even over longer time intervals (Ventura-Bort et al., 2016a, 2016b; Weymar et al., 2013) when associated with aversive context conditions, demonstrating a spill-over effect of an arousing context to neutral stimuli. This effect was absent in the three studies using highly socially relevant cues (i.e., faces). The faces were explicitly brought to attention to the participants. This was accomplished either by a passive instruction to attentively view all presented faces (Study 1 and 3), by giving an explicit learning instruction for a later recognition test (Study 2), or by applying a visual working memory task in which changes of face identity had to be detected (Study 2). However, the arousing contexts occupied the attentional resources which is likely a contributing factor to the impaired face recognition performance in all studies. Furthermore, faces competed for attention with the salient aversive context conditions in Study 2 through explicit learning instruction. Performance was then impaired during threat and safety improved memory as no conflict of resources was present during safety, adding to an attentional account of the findings. Additionally, compensatory strategies seemed efficient enough to improve accuracy under threat in working memory when arousal was task-irrelevant (ACT; Robinson et al., 2013; Eysenck et al., 2007). Taken together, this suggests that arousal as a context condition does not impair short-term memory when sufficient cognitive resources are available and arousal does not act as a distracting element in the task flow (Stout et al., 2013; Sessa et al., 2011). However, arousing contexts bind attention and impair the building of associations to closely attached cues when this is not enforced externally (e.g., through instruction), thus reducing memory for the cues. Therefore, the studies add valuable insight into the complex interaction of arousal, attention, and memory.

On a more technical level, especially Study 1 served to test an adapted and well-established statistical model (hierarchical MPT; Bayen et al., 1996; Batchelder & Riefer, 1990) to quantify memory performance that may be used in follow-up studies of this thesis. This model made it possible to disentangle the different underlying latent cognitive mechanisms that contribute to performance and minimize confounders in the form of guessing biases. With this, item and source memory as well as the influence of arousal could be assessed independently. In addition, the use of a hierarchical extension of the model provided the opportunity to calculate possible associations between parameter estimates, such as guessing that a face was old (i.e., parameter b), and individual differences, such as social anxiety or ACE scores. This may prove particularly helpful in the future to investigate memory in clinical populations more rigorously (Arnold, Bayen, & Böhm, 2015).

In contrast to the behavioral findings, contextual factors modulated person perception on the neural level (i.e., ERPs) even in the absence of conscious recognition (Diana et al., 2011; Addante et al., 2012). This dissociation between perceptual and behavioral processing could be replicated across the use of different social fear learning pathways and in healthy and traumatized individuals. Both, differential processing due to characteristics of the faces (i.e., old/new effect, cognitive load effect on N170) as well as based on (former) threatening and safe context (i.e., LPP, CDA) were found in Study 1 and in Study 2 for both memory tasks. This supports previous findings of advanced sensitivity of neural processing over behavior during subtle context differentiation as well as a restrained access to working memory capacity (Sessa et al., 2011; Luck et al., 1996; Stout et al., 2013). This observed dissociation is assumed to incorporate processing modulations that underlie a cognitive task but does not translate into overt behavioral changes (Wilkinson & Halligan, 2004).

Interestingly, this dissociation was not observed for the activation of psychophysiological defense mechanisms in Study 3. Previously threat-associated faces did not trigger a potentiation in somatic and automatic responding (i.e., startle EMG, EDA), when threat had to be inferred and was not actually present (i.e., during recognition). Exactly as in Studies 1 and 2, this experimental phase was characterized by a complete absence of the possibility of threatening stimulation by removing the shock electrode. Thus, the activation of the somatic and automatic nervous system seems more strongly dependent on external cues in order to be modulated by aversive arousal when conscious recognition – or at least a sense of familiarity - of threatening events is lacking (Diana et al., 2011). It may therefore be less sensitive to the priming of former perceptual differences compared to neural processing when awareness of these differential conditions is lacking.

These findings do not support the notion that perceptual biases necessarily translate into physiological arousal. Furthermore, neither perceptual bias nor defense behavior directly relate to threat-associated identity recognition benefits or deficits. Directed attentional mechanisms during learning, as described above, are one possible explanation, with memory and physiological processing being more vulnerable to a lack of attention compared to perceptual processing. Furthermore, the formation of a memory trace requires time to consolidate (Squire et al., 2015; Ventura-Bort et al., 2016b), which in turn should increase the ability to recognize the stimuli and then reflect processing biases. Lastly, the complexity and difficulty of the tasks (recognizing unfamiliar faces, change detection of faces during high load), could have disguised the effect of arousal on the behavioral level. Future studies should therefore account for attentional mechanisms and include an easier task with more time to consolidate.

Influence of anxiety and adverse childhood experiences. On the behavioral level, anxiety had an impact on the association of arousal and memory only when the conditions for successful task performance were more adequately supplied (Study 2). While anxiety did not modulate the findings of context independent low face recognition (Study 1 and 3), the facilitating effect of safety for face recognition seemed to increase by higher levels of social anxiety (Study 2). In contrast, individuals with high social and state anxiety appeared to perform better under threatening contextual conditions during working memory (Study 2).

The literature regarding memory performance in pathological anxiety- and trauma-related processing is mixed. There are indications of a generally arousal-related impaired functioning in working memory with an intact long-term memory (Robinson et al., 2013; Dolcos et al., 2020; Friedman & Johnson, 2000). A possible explanation may be the adaptive advantage of (long-term) memorizing of threatening information, thus being able to recognize danger in future situations and the disadvantage of a threat-bias in information processing when it interferes with performance success (i.e., working memory). However, there are also indications of opposite effects of arousal, with studies showing impaired long-term memory and intact working memory processing (Robinson et al., 2013; Sauro et al., 2003; Stout et al., 2013). This highlights the role of different components determining the impact of arousal on memory. For instance, different types of ACE can have different effects on neural development, behavior, and cognitive functioning (Guinosso et al., 2016; Sheridan & McLaughlin, 2014; Herzog & Schmahl, 2019). It is suggested that childhood experiences of neglect result in a deprived environment lacking social, cognitive, and emotional inputs. This in turn negatively affects the development of the cortex and leads to impairments in tasks, that depend on this area, for instance complex cognitive tasks, executive functioning, spatial navigation. Abuse has a greater impact on emotional processing and learning and may be more related to changes in the hippocampus and amygdala.

Individuals with adverse childhood experiences showed overall intact memory processing regardless of the severity of ACE and comorbid psychopathology such as anxiety, with no overt behaviorally biased or generalized threat processing. Exploratory results from ACE subtype analyses indicated that individuals with a history of childhood abuse (emotional and sexual) performed better and benefitted more form threatening context conditions both in recognizing faces and during working memory. In contrast, while emotional neglect was also associated with enhanced face recognition from a threatening context, working memory performance was enhanced under safe conditions (Robinson et al., 2013). This highlights a heightened vigilance in arousing situations. The beneficial or detrimental effect of arousal seems, however, to depend on the type of traumatic experience. Abusive experiences inherently require a state of constant alertness and the need to scan the environment for potential danger cues, whereas neglect may be characterized mainly by the feeling of absence of safety. This could have led to the benefit of threat in abuse and the more differentiated effect of threat in neglect, with enhanced working memory once a state of safety is reached. Future research should incorporate a more differentiated design, including task (non)related threatening and neutral distracting information. Distraction and inhibition ability seem to largely contribute to memory processes, especially in relation to anxiety and trauma (Stout et al., 2013; Ward et al., 2020; MacNamara et al., 2011).

Regarding neural processing, a modulating effect of social anxiety on the differential effect of threat and safety was revealed, both during encoding and recognition. The perceptual difference in threat- and safety-face compounds was elevated for socially anxious individuals during encoding (Study 1). Arousal-based psychophysiological responding (i.e., threat potentiated SCR) also increased with enhanced levels of anxiety and depression (Study 3). Furthermore, even when not consciously retrievable, the pronounced perceptual processing of threatening cues was observed in more socially anxious participants during recognition (Study 1). The differentiation in defense behavior, however, did not persevere into the recognition phase and was thus not modulated by anxiety, probably due to habituation effects. Thus, anxiety modulates perceptual bias but not defensive responding towards threat during aversive anticipation and thereupon face perception (Wieser et al., 2014). At low-moderate levels of anxiety, the translation of perceptual bias into defense activation seems to lack significant impact (Mitte, 2008).

These effects were not modulated by subtypes of ACE. However, the threat-elevated ERP effect during recognition was generally reversed in individuals with a history of childhood maltreatment compared to healthy individuals (Studies 1 & 2). While an enhanced negative-going waveform for former threat compared to safe and new faces was found for the healthy sample, individuals with ACE showed an enhanced LPP to threat compared to safety. This could be an indication of an enhanced threat bias in traumatized individuals with facilitated and motivational driven processing of threatening information (Schupp et al., 2004). The focus on potentially harming and dangerous situations and cues seems adaptive from an evolutionary perspective, especially after experiences of severe maltreatment. However, when this persists into safe contexts, this can become maladaptive and lead to pathological fear.

5.2 Constraints on Generalizability and Future Directions

The presented studies share significant contributions regarding the impact of arousal on learning and memory processes and the modulating role of anxiety and childhood trauma experiences using a broad range of (psychophysiological) methods to gain comprehensive insights (self-report, behavioral measures, EEG/ERP, startle EMG, EDA). However, there are also some factors that limit the meaningfulness and interpretability of the discussed results. These factors can directly translate into future research implications and the design of future studies.

Improving the binding of arousal and faces. Regarding both memory tasks, the key interest lay in the behavior and perception towards the contextual conditions of the short- and long-term memory task. The context information was therefore more important than the depicted faces themselves, both in terms of result evaluation and perceived significance for physical integrity (i.e., bodily harm by electrical shock). However, for both tasks the identification of the facial stimuli was essential for successful task performance (Arnold et al., 2021). Identifying the faces was the prerequisite for any accurate source identification in the item/source memory task and the focus in the visual working memory task in order to detect changes, with context being a negligible factor. This could have led to contrasting effects of attention allocation in Study 1 compared to Study 2. While face encoding was not enforced in

Study 1, participants focus may have been most prominently on the context conditions disregarding the (less relevant) faces, which could in turn have added to the interpretation of poor face memory. In Study 2, however, this imbalance between item and context may have led to a disregard of the contextual settings during the memory task as the context became task-irrelevant through the explicit learning instruction for the faces. Especially the presence of a negative element (i.e., possibility of receiving an electrical shock) has been shown to decrease associations between items and their surroundings (Bisby, et al., 2016).

On a neural level, the arousing element may be involved in the downregulation of the hippocampus, which plays an important role in building associations between elements while simultaneously up-regulating the amygdala, thus facilitating the encoding of the arousing information (Bisby et al., 2018; Damasio, 1989; Davachi, 2006; Eichenbaum et al., 2007; O'Keefe & Nadel, 1978; Jacobs & Nadel, 1998; Bisby & Burgess, 2017). Nevertheless, attention accounts are not sufficient to explain the lack of associations between faces and the (non)arousing contexts. In line with previous findings, the differences in neural perception and threat-potentiated activation of the somatic and autonomic nervous system are signs that attention was paid towards the threatening backgrounds, regardless of experimental or study design (Schellhaas et al., 2020; Bublatzky et al., 2022; Bisby & Burgess, 2014). Consequently, measures of attentional focus (e.g., eye-tracking) could be beneficial in disentangling the effect of directed attention and formed association between an arousing and neutral event/stimulus in future studies.

Strengthening the association between contexts and faces was attempted from Study 1 to Study 2. Increasing the viewing time of the face stimuli and organizing the alternating threat and safety contexts in a block-wise rather than a (random) trial-wise fashion together with the explicit instruction to memorize the faces improved face recognition. These changes in experimental design were assumed to lead to less distraction due to an everchanging environment and to the adaptation of slower defense responding to the current (non-)arousing circumstances.

However, replicating findings of impaired source recognition in relation with facial stimuli and threat (Arnold et al., 2021), successful source attribution failed in all three studies. In Study 3, face recognition was deliberately held low to address the question whether this perceptual difference for "non"-recognized faces would translate directly into defense mechanisms (startle EMG, EDA). This assumption was not supported by the results. Therefore, it remains unclear what drove the (reverse) findings of dissociation between the inability to actively retrieve associated source information and perceptual processing that differed in regard to the initial (non)arousing contexts in the first two studies: the inability to consciously retrieve the item and source information as a result of the inaccessibility of an existing memory trace or actual forgetting.

Hence, memory performance should be boosted further with the aim of enhancing source identification and strengthening the association of the item and source information. The central facial information should not only have a very close temporal and spatial connection to the contextual information but should be incorporated semantically in the perception of threat and safety. This could be done, for instance, by framing a semantic meaning to the faces inside a threatening or safe situation or by explicitly stressing the threatening or safe nature of the encounters (Hennings et al., 2021; Dunsmoor et al., 2015). Moreover, similar learning and recognition conditions are suggested to improve both stages of memory processing (Tulving & Thomson, 1973; Murnane et al., 1999). Thus, memory is context-dependent with better memory for similar encoding and recall conditions (Godden & Baddeley, 1975), including physiological and emotional states (e.g., state-dependent and mood-dependent learning; Clark et al., 1983). Consequently, the associative arousing information could have been present throughout the memory tasks, including the recognition session in all three studies. Enhancing the comparability of the two memory tasks, the procedure of including arousal in both, encoding and recognition, would assimilate to the settings of the visual working memory task, in which threat and safety were a constant alternating feature. This could have been accomplished by presenting the faces in both context conditions with a forced-choice selection during recognition. Secondly, by triggering familiarity-based retrieval processes, associative memory could be improved (Rugg & Curran, 2004; Diana, et al., 2011; Kahn et al., 2004). Faces of known, popular persons intermixed with new and unfamiliar faces are suited to disentangle knowing from guessing mechanisms. This would additionally tackle the finding of hindered remembering of unfamiliar faces in general (Burton & Jenkins, 2011). Lastly, the use of novel, neutral items other than faces could reveal a possible generalization of the effects over a broad range of situations and stimuli.

Attentional focus, learning and memory consolidation. New information requires attentional focus and time to consolidate (McGaugh, 2000; Ventura-Bort et al., 2016b) and the mnemonic benefit for arousal gets stronger with longer retention delays (Kensinger, 2009). Therefore, longer consolidation periods between the encoding and recognition phase seem pertinent to boost item and source memory (e.g., a gap of one hour, day or even week; Ventura-Bort et al., 2016b). This could also dampen the occurrence of task demands, which were especially high during study 2, where participants performed both memory tasks over a relatively short period of time, with only a little break in between. Due to the nature of the task and the necessity of a high number of trials for the analyses of the EEG/ERP data, the visual working memory task (change detection) was especially strenuous for the participants. This could have led to a drop in motivational focus and attention, apparent in a performance drop during the high load condition behaviorally, which was not mirrored in neural processing measured by the CDA. Increasing the attentional focus on the items and their contexts in the to be performed task and performance-based incentives are needed to rule out alternative explanations of the findings that depend on motivational factors in task performance.

Different learning instructions were used for the three studies, limiting the comparability between them. Study 1 used an incidental learning design by instructing the participants to view all presented pictures passively but attentively in front of their respective backgrounds without making it explicitly aware to them that a test will occur. In order to enhance face recognition in Study 2, an intentional learning instruction was used for the faces but not for the contexts. The participants were aware of a memory test for the faces but not for the contexts, which was then a surprising element during the recognition test. As the aim of study 3 was to tackle underlying defense mechanisms of unrecognized but formerly dangerous situations, a return to the incidental learning instruction was chosen to keep face recognition low. However, the enhanced presentation time for the faces from Study 1 to Study 2 (from 1s to 6s) was retained due to the slower signaling of the EMG and EDA responses, that were the key measures in Study 3. Intentional learning as well as prolonged presentation time was used to improve the retrieval of learnt stimuli, rendering it difficult to disentangle the influence of either of these changes that were made between the studies and to interpret and explain the changes in behavioral and neural findings. Therefore, a more exhaustive design including every combination of the chosen task features (short vs. long presentation time; explicit vs. incidental learning of the faces) would be needed in future studies.

Cognitive functioning in everyday life. Implications of impaired cognitive functioning in psychological experiments in the laboratory can have little ecological validity and the interpretability of the findings for real-world functioning is somewhat impeded. Thus, a measure of the basic level of cognitive functioning is lacking and it remains unclear whether the findings are solely based on the task properties or if the individuals of the studies are self-selectively high- or low-cognitive-functioning. Individuals with a broad range of psychopathology following trauma report cognitive impairments in their everyday life. Such impairments often not only concern higher cognitive functioning (e.g., problem solving and decision making), but also low demanding tasks (Buckholtz & Meyer-Lindenberg, 2012; Snyder et al., 2015; Bloemen, et al., 2018). Even between and within healthy individuals cognitive functioning varies, depending on various factors such as age and other individual differences. Therefore, a measure of cognitive functioning in everyday life would contribute to clarify whether the effects of arousal on memory in the three studies are associated to general cognitive impairments. This would enhance interpretability and generalizability of the findings especially in Study 2, in which impairments in cognitive functioning due to childhood maltreatment are assumed.

For this purpose, a useful concept is cognitive failures, one measure for which is the Cognitive Failures Questionnaire (Broadbent et al., 1982; Wallace et al., 2002; Klumb, 1995). Self-reported data acquired from diaries can also be used (e.g., Unsworth et al., 2012). Cognitive failure refers to the lack of success with simple, cognitive-based tasks, that can be carried out without making errors (e.g., remember to carry the grocery list to the grocery store). The discrepancies between reported cognitive failures in everyday life and those measured in a laboratory might reflect subjective worry about one's own state of cognition when other (state- and trait-related) factors are involved. In contrast, the stakes are low in a rather sterile laboratory environment where objective measures are used and cognitive impairments carry less weight (Wilhelm, Witthöft, & Schipolowski, 2010; Carrigan, & Barkus, 2016). The incorporation of cognitive failures into future studies could improve the translational relevance of measures of ability to real-world functioning and heighten ecological validity (Carrigan & Barkus, 2016).

Fear learning pathways. A limitation common in all types of fear learning paradigms is the possibility of an extinguished CR during test, when the US is constantly omitted (Haaker et al., 2017). This seems less likely in this case, as differential effects of threat and safety were found even in the complete absence of contextual cues. However, a comparison to direct fear learning (i.e., conditioning), the actual experience of threat and safety, seems important to rule out any effects that could have modulated the findings of a dissociation between behavioral and neural findings.

A comparison of instructional and observational fear learning in study 2 showed that the different learning pathways did not differ with regard to a successful threat introduction. A well-established protocol for observational fear learning was used (Haaker et al., 2017), however, a confounding influence of possible additional verbal instruction for the observational learning paradigm cannot be ruled out. The instructions leading to the video with the learning model had to be rather specific, excluding only the information on the relevant threat stimulus, transmitted then by the video presentation. Although questions regarding threat and safety contingencies were not directly answered during the experiment, and the participants were referred to the video, disentangling a differential effect of observation and instruction needs more replicating studies. These should vary the degree of instructions given before participants view the observation videos and change features of the model and depicted observation scene.

Furthermore, distinguishable effects of different learning pathways are mostly evident in differential information flow between neural networks (Lindström et al., 2018; Olsson & Phelps, 2007). For fear conditioning, the amygdala, sensory cortices, thalamus and regions associated with pain processing, ACC, and insular cortex have been shown to be involved in transmission of threat. For observational learning, the cortical representation of empathetic pain may modulate threat transfer via ACC and insular cortex (Olsson & Phelps, 2004). For instructional learning, the involvement of a cortical network bound by the hippocampus in threat association seems most likely (Olson & Phelps, 2004). The inclusion of an additional neuroimaging methods, e.g. functional magnetic resonance imaging (fMRI), seems pertinent. This would provide complimentary data with high source location in addition to the high temporal resolution data from EEG, bringing the time course of memory processing together with its source origins. Neural components and neuronal networks could be identified that are engaged in the acquisition and expression of fear through direct and indirect learning pathways and that are involved in other relevant processes as well, for instance in different memory processes such as source memory (Mitchell & Johnson, 2009), or individual differences based on psychopathology and anxiety.

Further biological markers could also be included as indices of stress and arousal, elicited by the aversive context. Using cortisol levels by collecting saliva samples, measuring endocrinological responses, would allow to test whether arousal leads to an increased release of cortisol which in turn is suggested to block memory retrieval (Wolf, 2009; Buchanan et al., 2006) or to facilitate memory, especially for emotional stimuli (Abercrombie et al., 2006; Buchanan & Lovallo, 2001).

5.3 Clinical Relevance

The findings discussed in this thesis result from strictly controlled laboratory settings and relied on community samples of participants with low-moderate levels of anxiety. Following a trans-diagnostic approach using a dimensional measure of childhood trauma, including five different subtypes (emotional and physical neglect and abuse, sexual abuse), a heterogeneous sample of participants resulted, reflecting the general population. In this sample, individuals often experienced different types of maltreatment simultaneously and exhibited mild to severe levels of ACE. While the derivation of clinical relevance from these findings must be treated cautiously as replication and application to psychopathological malfunctioning is needed, there are several clinical implications resulting from the use of this highly ecological sample.

There are indications that different types and timings of adverse childhood experiences affect behavior, psychopathology, cognitive functioning, and neural development differently, with the occurrence of multiple adversities having the most detrimental effect (Sheridan & McLaughlin, 2014; Herzog & Schmahl, 2018; Guinosso et al., 2016). There were no indications of impaired short- or long-term memory functioning in the individuals with ACE in this thesis. Individuals in general, including those without ACE, had difficulties in recognizing unfamiliar face identities and changes thereof and were incapable of reconstructing the (non)arousing context conditions of former encounters. Nevertheless, perceptual processing differed between former threat- and safe-face compounds, indicating that conscious recognition is not mandatory for sensitizing the perceptual system through arousal. Moreover, this effect was reversed in the individuals with ACE. Together with the findings of potentiated threat processing in (social) anxiety and depression as well as enhanced memory performance under threat for sexually abused, emotionally abused and neglected individuals, this strongly indicates a behavioral and perceptual bias for threat (Bar-Haim et al., 2007; Pergamin-Hight et al., 2015; Wald et al., 2013; Friedman & Johnson, 2000). This does not seem to necessarily negatively impact memory, even if arousal is task-irrelevant, as it was the case in the presented studies. On the contrary, this focus can have beneficial effects on memory and should not be considered detrimental per se.

Furthermore, recent research implicates that deficient safety learning is a key element of anxiety- and trauma-related psychopathology rather than threat-biases (Christianson et al., 2012; Grupe & Nitschke, 2013; Jovanovic et al., 2012; van Rooij & Jovanovic, 2019; Laing & Harrison, 2021). Impaired safety learning could therefore be the key to cognitive impairments. Patients and in general individuals suffering from maladaptive anxiety and traumatic experiences would benefit from interventions, learning and training programs that target safety learning and establish safety cues in aversive contexts that not only indicate the absence of danger but actively promote the presence of safety.

Lastly, memory functioning in the studies of this thesis was straightforward, with no distracting elements present or the necessity to actively inhibit the processing of task-irrelevant elements. Individuals with a history of trauma and anxiety often report impaired cognitive functioning in everyday life, mostly including problems concentrating. Increased ecological conditions, reflecting the complex environment, and containing multiple cues that interact with each other and serve as distractors are needed to deepen the understanding on how the lack of findings of impaired memory functioning translates into real-world cognitive functioning.

Taken together, individuals with adverse childhood experiences and anxiety did not exhibit impaired memory functioning, evident in intact face processing and recognition. There are indications of a behavioral and perceptual bias towards threat that can be compensated in terms of successful performance. Deficits in safety learning should be incorporated into future research designs to further investigate the reported difficulties in real-life cognitive functioning.

Summary

Recognizing and adequately responding to dynamic environmental conditions is crucial for adaptive and successful behavior. Threatening- and anxiety-inducing situations require well-functioning cognitive, defensive, and neural processes to prevent harm, especially when they are ambiguous. When recognized, threat cues can evoke general arousal and emotions such as fear and corresponding actions can be initiated (e.g., avoidance). The memory-enhancing effects of emotionally arousing stimuli are well documented. Under certain circumstances, the beneficial effect of arousal can spill over from an arousing context to otherwise neutral stimuli and enhance memory of these events. This spillover effect seems to depend on the memory modality and inter-individual differences such as anxiety or chronic stress exposure from childhood trauma, such as adverse childhood experiences (ACE). ACE are often associated with stress and anxiety-related disorders in adulthood. Symptom severity is linked to learning and memory deficits, which could connect childhood maltreatment and psychopathology. How childhood trauma might affect memory distortions in emotionally arousing settings is not well-understood.

The aim of this thesis is to examine the differential effect of threat and safety on neutral face perception, short-term processing (i.e., working memory) and long-term storage (i.e., recognition). Additionally this thesis focuses on the impact of adverse childhood experiences and anxiety on memory processing under threat. Threat was introduced using social fear learning paradigms, increasing the social nature of the tasks and accounting for vicariously and verbally acquired anxiety in psychopathology. Key metrics include measures of recognition and memory performance (Studies 1-3), event-related brain potentials (ERP in EEG; Studies 1 & 2) and measures of the somatic and autonomic nervous systems as indicators of defensive behaviors (startle Electromyography [EMG]; Electrodermal Activity [EDA]; Study 3).

Threat expectation was successfully induced via social learning and was demonstrated in all three studies (verbal threat learning in Study 1-3, observational fear learning in Study 2). This was true for both healthy participants and participants with ACE and varying levels of anxiety, suggesting intact threat learning and perception in trauma-predisposed individuals. Threat-associated faces also led to increased neural processing and psychophysiological responses. This arousal-based attention was evidenced by enhanced early parietal-occipital and late fronto-central negative potentials in EEG as well as threat-potentiated defensive behavior activation (startle reflex EMG, EDA), higher evaluations of perceived threat and arousal and lower evaluation of valence.

During recognition, all participants had difficulty identifying previously seen faces from new faces. They were even less able to recall the associated contexts (Study 1-3). Prolonging the viewing time during encoding phases, slightly improved person identification (Study 2 & 3). Learning instructions triggered increased recognition of faces from safe contexts in low anxious individuals with ACE (Study 2). In contrast, highly anxious participants seemed to benefit from threatening contexts regarding short-term face processing. Overall, short-term memory was affected neither by ACE nor by differing context conditions, but it was impaired by high cognitive strain. Intriguingly, however, brain activity did differentiate previously seen faces from newly presented pictures (old/new ERP/EEG effect; Study 1 & 2).

Most importantly, we found evidence of a neural differentiation between former threat-associated faces and both safe and new faces (Study 1 & 2). In Study 1, distributed late negativities emerged for threat-associated faces. This effect was reversed in the ACE sample which exhibited pronounced late positive potential for threat-associated faces relative to safety faces (Study 2). In working memory, more neural processing resources were allocated towards threat especially during high cognitive load conditions (Study 2). These findings support the view that contextually threatening conditions critically modulate face perception even in the absence of conscious recognition. They also point to a decreased capacity for faces when cognitive resources are limited during potentially harmful situations. This could serve as a neural marker for a threat bias in traumatized individuals.

Study 3 showed that conscious memory of faces associated with past dangerous situations appears to be necessary to better prepare individuals for future, and potentially equally threatening encounters. The dissociation in perceptual processing and overt (failed) recognition behavior did not translate into the activation of defensive systems (i.e., startle EMG, EDA) in the absence of actual threat and the associated memory. These findings do not support the notion that perceptual biases and physiological arousal directly relate to threat-associated identity recognition deficits.

Taken together, the three empirical studies included in this thesis show that arousal impeded face identification only when the circumstances of recognizing faces formerly encountered in threat or safety contexts were particularly favorable. Conscious awareness of threat seems required to trigger defensive behavior but not needed to trigger differential neural processing. This suggests that perceptual bias, overt behavior, and cognitive functioning are dissociated. A similar but opposite threat bias was present for healthy individuals and individuals with ACE, indicating an allocation of attention towards arousing stimuli in individuals who experienced childhood trauma. These findings suggest potential for clinical application of reducing threat expectations and inducing safety expectations through social learning. However, future research is needed to disentangle the effects of trauma and anxiety on memory processes.

References

Abercrombie, H. C., Speck, N. S., & Monticelli, R. M. (2006). Endogenous cortisol elevations are related to memory facilitation only in individuals who are emotionally aroused. *Psychoneuroendocrinology, 31*(2), 187-196. https://doi.org/10.1016/j.psyneuen.2005.06.008

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *Neuroimage, 62*(1), 439-450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current opinion in neurobiology, 12*(2), 169-177. https://doi.org/10.1016/S0959-4388(02)00301-X

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.).

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anderson, A. K., Wais, P. E., & Gabrieli, J. D. (2006). Emotion enhances remembrance of neutral events past. *Proceedings of the National Academy of Sciences, 103*(5), 1599-1604. https://doi.org/10.1073/pnas.0506308103

Arnold, N. R., Bayen, U. J., & Böhm, M. F. (2015). Is prospective memory related to depression and anxiety? A hierarchical MPT modelling approach. *Memory, 23*(8), 1215–1228. ttps://doi.org/10.1080/09658211.2014.969276

Arnold, N. R., Bayen, U. J., & Smith, R. E. (2015). Hierarchical multinomial modeling approaches. *Experimental psychology, 62,* 143-152. https://doi.org/10.1027/1618-3169/a000287

Arnold, N. R., Bayen, U. J., Kuhlmann, B. G., & Vaterrodt, B. (2013). Hierarchical modeling of contingency-based source monitoring: A test of the probability-matching account. *Psychonomic Bulletin & Review, 20*(2), 326–333. https://doi.org/10.3758/s13423-012-0342-7

Arnold, N. R., González Cruz, H., Schellhaas, S., & Bublatzky, F. (2021). A multinomial modelling approach to face identity recognition during instructed threat. *Cognition and Emotion, 35*(7), 1302–1319. https://doi.org/10.1080/02699931.2021.1951175

Askew, C., & Field, A. P. (2007). Vicarious learning and the development of fears in childhood. *Behaviour Research and Therapy, 45*(11), 2616-2627. https://doi.org/10.1016/j.brat.2007.06.008

Askew, C., & Field, A. P. (2008). The vicarious learning pathway to fear 40 years on. *Clinical psychology review, 28*(7), 1249-1265. https://doi.org/10.1016/j.cpr.2008.05.003

Atkinson, R. C., & Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*. In Psychology of learning and motivation (Vol. 2, pp. 89-195). Academic Press.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience, 139*(1), 201-208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Awipi, T., & Davachi, L. (2008). Content-specific source encoding in the human medial temporal lobe. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 769. https://doi.org/10.1037/0278-7393.34.4.769

Aylward, J., Valton, V., Goer, F., Mkrtchian, A., Lally, N., Peters, S., Limbachya, T., & Robinson, O. J. (2017). The impact of induced anxiety on affective response inhibition. *Royal Society Open Science, 4*(6), 170084. https://doi.org/10.1098/rsos.170084

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

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews neuroscience, 4*(10), 829-839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OuP Oxford.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Routledge.

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice Hall: Englewood cliffs.

Barch, D. M., & Smith, E. (2008). The cognitive neuroscience of working memory: relevance to CNTRICS and schizophrenia. *Biological psychiatry, 64*(1), 11-17. https://doi.org/10.1016/j.biopsych.2008.03.003

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M.J., & Van Ijzendoorn, M.H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological bulletin, 133*(1), 1. https://psycnet.apa.org/doi/10.1037/0033-2909.133.1.1

Barnacle, G. E., Tsivilis, D., Schaefer, A., & Talmi, D. (2018). Local context influences memory for emotional stimuli but not electrophysiological markers of emotion-dependent attention. *Psychophysiology, 55*(4), e13014. https://doi.org/10.1111/psyp.13014

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological review, 97*(4), 548. https://doi.org/10.1037/0033-295X.97.4.548

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*(1), 57-86. https://doi.org/10.3758/BF03210812

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 197-215. https://doi.org/10.1037/0278-7393.22.1.197

Beck, A. T., & Clark, D. A. (1997). An information processing model of anxiety: Automatic and strategic processes. *Behaviour research and therapy, 35*(1), 49-58. https://doi.org/10.1016/S0005-7967(96)00069-1

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature neuroscience, 4*(6), 645. https://doi.org/10.1038/88477

Bell, R., & Buchner, A. (2010). Valence modulates source memory for faces. *Memory & Cognition, 38*(1), 29-41. https://doi.org/10.3758/MC.38.1.29

Bell, R., & Buchner, A. (2011). Source memory for faces is determined by their emotional evaluation. *Emotion, 11*(2), 249-261. https://doi.org/10.1037/a0022597

Bell, R., Mieth, L., & Buchner, A. (2017). Emotional memory: No source memory without old–new recognition. *Emotion, 17*(1), 120. https://doi.org/10.1037/emo0000211

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex, 20*(12), 2874-2883. https://doi.org/10.1093/cercor/bhq038

Bergmann, H. C., Rijpkema, M., Fernandez, G., & Kessels, R. P. (2012). The effects of valence and arousal on associative working memory and long-term memory. *PLoS One, 7*(12), e52616. https://doi.org/10.1371/journal.pone.0052616

Bernstein, D.P., Fink, L., Handelsman, L., & Foote, J. (1998). *Childhood trauma questionnaire. Assessment of family violence: A handbook for researchers and practitioners*. https://psycnet.apa.org/doi/10.1037/t02080-000

Bisby, J. A., & Burgess, N. (2014). Negative affect impairs associative memory but not item memory. *Learning & Memory, 21*(1), 21–27. https://doi.org/10.1101/lm.032409.113

Bisby, J. A., & Burgess, N. (2017). Differential effects of negative emotion on memory for items and associations, and their relationship to intrusive imagery. *Current Opinion in Behavioral Sciences, 17*, 124-132. https://doi.org/10.1016/j.cobeha.2017.07.012

Bisby, J.A., Horner, A.J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology: General, 147*(2), 243. http://dx.doi.org/10.1037/xge0000356

Bisby, J. A., Horner, A. J., Hørlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience*, *11*(6), 981-990. https://doi.org/10.1093/scan/nsw028

Bisson, J. I., Roberts, N. P., Andrew, M., Cooper, R., & Lewis, C. (2013). Psychological therapies for chronic post-traumatic stress disorder (PTSD) in adults. *Cochrane database of systematic reviews, (12*). https://doi.org/10.1002/14651858.CD003388.pub4

Blair, K. S., Otero, M., Teng, C., Geraci, M., Lewis, E., Hollon, N., ... & Pine, D. S. (2016). Learning from other people's fear: amygdala-based social reference learning in social anxiety disorder. *Psychological medicine*, *46*(14), 2943-2953. https://doi.org/10.1017/S0033291716001537

Bloemen, A. J. P., Oldehinkel, A. J., Laceulle, O. M., Ormel, J., Rommelse, N. N. J., & Hartman, C. A. (2018). The association between executive functioning and psychopathology: general or specific?. *Psychological medicine*, *48*(11), 1787-1794. https://doi.org/10.1017/S0033291717003269

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*(1), 1-15. https://doi.org/10.1111/j.1469-8986.2005.00271.x

Böcker, K. B., Baas, J. M., Kenemans, J. L., & Verbaten, M. N. (2004). Differences in startle modulation during instructed threat and selective attention. *Biological Psychology*, *67*(3), 343-358. https://doi.org/10.1016/j.biopsycho.2004.01.001

Bounoua, N., Miglin, R., Spielberg, J. M., & Sadeh, N. (2020). Childhood assaultive trauma and physical aggression: links with cortical thickness in prefrontal and occipital cortices. *Neuroimage: clinical*, *27*, 102321. https://doi.org/10.1016/j.nicl.2020.102321

Boscarino, J. A. (2006). External-cause mortality after psychologic trauma: the effects of stress exposure and predisposition. *Comprehensive psychiatry, 47*(6), 503-514. https://doi.org/10.1016/j.comppsych.2006.02.006

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry, 25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, *18*(2), 379-390. https://doi.org/10.1037/0278-7393.18.2.379

Bradley, M. M., Moulder, B., & Lang, P. J. (2005). When good things go bad: the reflex physiology of defense. *Psychological science, 16*(6), 468-473. https://doi.org/10.1111/j.0956-7976.2005.01558.x

Bradley, M. M., Silakowski, T., & Lang, P. J. (2008). Fear of pain and defensive activation. *PAIN®, 137*(1), 156-163. https://doi.org/10.1016/j.pain.2007.08.027

Bradley, M. M., Zlatar, Z. Z., & Lang, P. J. (2018). Startle reflex modulation during threat of shock and “threat” of reward. *Psychophysiology, 55*(2), e12989. https://doi.org/10.1111/psyp.12989

Bremner, J.D., Vythilingam, M., Vermetten, E., Southwick, S.M., McGlashan, T., Staib, L.H., ... & Charney, D.S. (2003). Neural correlates of declarative memory for emotionally valenced words in women with posttraumatic stress disorder related to early childhood sexual abuse. *Biological psychiatry*, *53*(10), 879-889. https://doi.org/10.1016/S0006-3223(02)01891-7

Brewin, C. R. (2011). The nature and significance of memory disturbance in posttraumatic stress disorder. *Annual review of clinical psychology*, *7*, 203-227. https://doi.org/10.1146/annurev-clinpsy-032210-104544

Brewin, C. R., Huntley, Z., & Whalley, M. G. (2012). Source memory errors associated with reports of posttraumatic flashbacks: A proof of concept study. *Cognition*, *124*(2), 234-238. https://doi.org/10.1016/j.cognition.2012.05.002

Brewin, C. R., Kleiner, J. S., Vasterling, J. J., & Field, A. P. (2007). Memory for emotionally neutral information in posttraumatic stress disorder: A meta-analytic investigation. *Journal of abnormal psychology, 116*(3), 448. https://doi.org/10.1037/0021-843X.116.3.448

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*(1), 1-16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Bröder, A., & Meiser, T. (2007). Measuring source memory. *Zeitschrift für Psychologie/Journal of Psychology, 215*(1), 52-60. https://doi.org/10.1027/0044-3409.215.1.52

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73*(1), 105-116. https://doi.org/10.1111/j.2044-8295.1982.tb01795.x

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied, 7*(3), 207. https://doi.org/10.1037/1076-898X.7.3.207

Bublatzky, F., & Schupp, H. T. (2012). Pictures cueing threat: brain dynamics in viewing explicitly instructed danger cues. *Social cognitive and affective neuroscience, 7*(6), 611-622. https://doi.org/10.1093/scan/nsr032

Bublatzky, F., Flaisch, T., Stockburger, J., Schmälzle, R., & Schupp, H. T. (2010). The interaction of anticipatory anxiety and emotional picture processing: An event-related brain potential study. *Psychophysiology, 47*(4), 687-696. https://doi.org/10.1111/j.1469-8986.2010.00966.x

Bublatzky, F., Gerdes, A., White, A. J., Riemer, M., & Alpers, G. W. (2014). Social and emotional relevance in face processing: happy faces of future interaction partners enhance the late positive potential. *Frontiers in human neuroscience, 8*, 493. https://doi.org/10.3389/fnhum.2014.00493

Bublatzky, F., Gerdes, A.B., & Alpers, G.W. (2014). The persistence of socially instructed threat: Two threat-of-shock studies. *Psychophysiology, 51*(10), 1005-1014. https://doi.org/10.1111/psyp.12251

Bublatzky, F., Guerra, P. M., Pastor, M. C., Schupp, H. T., & Vila, J. (2013). Additive effects of threat-of-shock and picture valence on startle reflex modulation. *PloS ONE, 8*, e54003. https://doi.org/10.1371/journal.pone.0054003

Bublatzky, F., Guerra, P., & Alpers, G. W. (2018). Verbal instructions override the meaning of facial expressions. *Scientific reports, 8*(1), 1-11. https://doi.org/10.1038/s41598-018-33269-2

Bublatzky, F., Guerra, P., & Alpers, G.W. (2020). Watch out, he's dangerous! Electrocortical indicators of selective visual attention to allegedly threatening persons. *Cortex, 131*, 164-178. https://doi.org/10.1016/j.cortex.2020.07.009

Bublatzky, F., Kavcıoğlu, F., Guerra, P., Doll, S., & Junghöfer, M. (2020). Contextual information resolves uncertainty about ambiguous facial emotions: behavioral and magnetoencephalographic correlates. *NeuroImage*, 116814. https://doi.org/10.1016/j.neuroimage.2020.116814

Bublatzky, F., Pittig, A., Schupp, H. T., & Alpers, G. W. (2017). Face-to-face: Perceived personal relevance amplifies face processing. *Social cognitive and affective neuroscience, 12*(5), 811-822. https://doi.org/10.1093/scan/nsx001

Bublatzky, F., Riemer, M., & Guerra, P. (2019). Reversing threat to safety: incongruence of facial emotions and instructed threat modulates conscious perception but not physiological responding. *Frontiers in psychology, 10,* 2091. https://doi.org/10.3389/fpsyg.2019.02091

Bublatzky, F., Schellhaas, S., & Guerra, P. (2022). The mere sight of loved ones does not inhibit psychophysiological defense mechanisms when threatened. *Scientific reports, 12*(1), 1-10. https://doi.org/10.1038/s41598-022-06514-y

Bublatzky, F., Schellhaas, S., & Paret, C. (under review). Comparing apples and oranges: Electrocortical correlates of decision-making and reward reversal learning under threat [Manuscript submitted for publication].

Buchanan, T. W., & Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology, 26*(3), 307-317. https://doi.org/10.1016/S0306-4530(00)00058-5

Buchanan, T. W., Tranel, D., & Adolphs, R. (2006). Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & memory, 13*(3), 382-387. https://doi.org/ 10.1101/lm.206306

Buckholtz, J. W., & Meyer-Lindenberg, A. (2012). Psychopathology and the human connectome: toward a transdiagnostic model of risk for mental illness. *Neuron, 74*(6), 990-1004. https://doi.org/10.1016/j.neuron.2012.06.002

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychological review, 112*(2), 291. https://doi.org/10.1037/0033-295X.112.2.291

Burton, A.M., & Jenkins, R. (2011). *Unfamiliar face perception*. The Oxford handbook of face perception, 28, 287-306.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248. https://doi.org/10.1111/1467-9280.00144

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and cognition, 4*(4), 410-421. https://doi.org/10.1006/ccog.1995.1048

Cahill, S. P., & Foa, E. B. (2007). *Psychological theories of PTSD.* Handbook of PTSD: Science and practice, 55-77.

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience, 14*(4), 1228-1258. https://doi.org/10.3758/s13415-014-0270-2

Carretié, L., Albert, J., López-Martín, S., & Tapia, M. (2009). Negative brain: an integrative review on the neural processes activated by unpleasant stimuli. *International Journal of Psychophysiology, 71*(1), 57-63. https://doi.org/10.1016/j.ijpsycho.2008.07.006

Carrigan, N., & Barkus, E. (2016). A systematic review of cognitive failures in daily life: Healthy populations. *Neuroscience & Biobehavioral Reviews, 63,* 29-42. https://doi.org/10.1016/j.neubiorev.2016.01.010

Chiu, Y. C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Frontiers in Psychology, 4,* 103. https://doi.org/10.3389/fpsyg.2013.00103

Christianson, J. P., Fernando, A. B., Kazama, A. M., Jovanovic, T., Ostroff, L. E., & Sangha, S. (2012). Inhibition of fear by learned safety signals: a mini-symposium review. *Journal of Neuroscience, 32*(41), 14118-14124. https://doi.org/10.1523/JNEUROSCI.3340-12.2012

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current opinion in neurobiology, 17*(2), 177-184. https://doi.org/10.1016/j.conb.2007.03.005

Clark, M. S., Milberg, S., & Ross, J. (1983). Arousal cues arousal-related material in memory: Implications for understanding effects of mood on memory. *Journal of Verbal Learning and Verbal Behavior, 22*(6), 633-649. https://doi.org/10.1016/S0022-5371(83)90375-4

Cocchini, G., Logie, R. H., Sala, S. D., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition, 30*(7), 1086-1095. https://doi.org/10.3758/BF03194326

Colloca, L., & Benedetti, F. (2009). Placebo analgesia induced by social observational learning. *PAIN®, 144*(1-2), 28-34. https://doi.org/10.1016/j.pain.2009.01.033

Cooper, E., Greve, A., & Henson, R. N. (2017). Assumptions behind scoring source versus item memory: Effects of age, hippocampal lesions and mild memory problems. *Cortex, 91*, 297-315. https://doi.org/10.1016/j.cortex.2017.01.001

Core Team, R. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/

Costa, V. D., Bradley, M. M., & Lang, P. J. (2015). From threat to safety: Instructed reversal of defensive reactions. *Psychophysiology, 52*(3), 325-332. https://doi.org/10.1111/psyp.12359

Craske, M. G., Wolitzky-Taylor, K. B., Mineka, S., Zinbarg, R., Waters, A. M., Vrshek-Schallhorn, S., Epstein, A., Naliboff, B., & Ornitz, E. (2012). Elevated responding to safe conditions as a specific risk factor for anxiety versus depressive disorders: evidence from a longitudinal investigation. *Journal of Abnormal Psychology, 121*(2), 315-324. https://doi.org/10.1037/a0025738

Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of vision, 10*(4), 16-16. https://doi.org/10.1167/10.4.16

Crowder, R. G. (1978). *Sensory memory systems*. In Perceptual Coding (pp. 343-373). Academic Press.

Dalgleish, T., Taghavi, R., Neshat-Doost, H., Moradi, A., Canterbury, R., & Yule, W. (2003). Patterns of processing bias for emotional information across clinical disorders: A comparison of attention, memory, and prospective cognition in children and adolescents with depression, generalized anxiety, and posttraumatic stress disorder. *Journal of clinical child and adolescent psychology, 32*(1), 10-21. https://doi.org/10.1207/S15374424JCCP3201_02

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural computation, 1*(1), 123-132. https://doi.org/10.1162/neco.1989.1.1.123

Danker, J. F., Hwang, G. M., Gauthier, L., Geller, A., Kahana, M. J., & Sekuler, R. (2008). Characterizing the ERP Old–New effect in a short-term memory task. *Psychophysiology*, *45*(5), 784-793. https://doi.org/10.1111/j.1469-8986.2008.00672.x

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current opinion in neurobiology*, *16*(6), 693-700. https://doi.org/10.1016/j.conb.2006.10.012

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual review of neuroscience, 15(*1), 353-375.

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology, 35*(1), 105-135. https://doi.org/10.1038/npp.2009.109

Davies, G., & Milne, A. (1982). Recognizing faces in and out of context. *Current psychology, 2*(1), 235-246. https://doi.org/10.1007/BF03186766

Debiec, J., & Olsson, A. (2017). Social fear learning: from animal models to human function. *Trends in cognitive sciences*, *21*(7), 546-555. https://doi.org/ 10.1016/j.tics.2017.04.010

Deco, G., & Rolls, E. T. (2005). Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *Journal of neurophysiology*, *94*(1), 295-313. https://doi.org/10.1152/jn.01095.2004

Deltomme, B., Mertens, G., Tibboel, H., & Braem, S. (2017). Instructed fear stimuli bias visual attention. *Acta Psychologica*. https://doi.org/10.1016/j.actpsy.2017.08.010

De Rosnay, M., Cooper, P. J., Tsigaras, N., & Murray, L. (2006). Transmission of social anxiety from mother to infant: An experimental study using a social referencing paradigm. *Behaviour research and therapy, 44*(8), 1165-1175. https://doi.org/10.1016/j.brat.2005.09.003

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1245-1255. https://doi.org/10.1098/rstb.1998.0280

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience, 18*(1), 193-222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Diana, R. A., Van den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, *1367*, 278-286. https://doi.org/10.1016/j.brainres.2010.10.030

Dimberg, U., & Öhman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion, 20*(2), 149-182. https://doi.org/10.1007/BF02253869

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. Emotion, 1(1), 5-11. https://doi.org/10.1037/1528-3542.1.1.5

Dolan, R. J. (2002). Emotion, Cognition, and Behavior. *Science, 298*(5596), 1191-1194. https://doi.org/10.1126/science.1076358

Dolan, R. J., & Vuilleumier, P. (2003). Amygdala automaticity in emotional processing. *Annals of the New York Academy of Sciences, 985*(1), 348-355. https://doi.org/10.1111/j.1749-6632.2003.tb07093.x

Dolcos, F., Denkova, E., & Dolcos, S. (2012). Neural correlates of emotional memories: A review of evidence from brain imaging studies. *Psychologia, 55*(2), 80–111. https://doi.org/10.2117/psysoc.2012.80

Dolcos, F., Katsumi, Y., Moore, M., Berggren, N., de Gelder, B., Derakshan, N., ... & Dolcos, S. (2020). Neural correlates of emotion-attention interactions: From perception, learning, and memory to social cognition, individual differences, and training interventions. *Neuroscience & Biobehavioral Reviews, 108*, 559-601. https://doi.org/10.1016/j.neubiorev.2019.08.017

Donchin, E. (1981). Surprise!… surprise?. *Psychophysiology, 18*(5), 493-513. https://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Dube, S.R., Felitti, V.J., Dong, M., Giles, W.H., & Anda, R.F. (2003). The impact of adverse childhood experiences on health problems: evidence from four birth cohorts dating back to 1900. *Preventive medicine, 37*(3), 268-277. https://doi.org/10.1016/S0091-7435(03)00123-3

Dunning, J. P., DelDonno, S., & Hajcak, G. (2013). The effects of contextual threat and anxiety on affective startle modulation. *Biological Psychology, 94*(1), 130–135. https://doi.org/10.1016/j.biopsycho.2013.05.013

Dunsmoor, J.E., Murty, V.P., Davachi, L., & Phelps, E.A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature, 520*(7547), 345-348. https://doi.org/10.1038/nature14106

Edwards, E.J., Edwards, M.S., & Lyvers, M. (2015). Cognitive trait anxiety, situational stress, and mental effort predict shifting efficiency: Implications for attentional control theory. *Emotion, 15*(3), 350. http://dx.doi.org/10.1037/emo0000051

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour research and therapy, 38*(4), 319-345. https://doi.org/10.1016/S0005-7967(99)00123-0

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci., 30*, 123-152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Eimer, M. (2000). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical neurophysiology, 111*(4), 694-705. https://doi.org/10.1016/S1388-2457(99)00285-0

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia, 45*(1), 15-31. https://doi.org/10.1016/j.neuropsychologia.2006.04.022

Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues* (Vol. 10). Ishk.

Ellis, A. W., Young, A. W., & Flude, B. M. (1990). Repetition priming and face processing: Priming occurs within the system that responds to the identity of a face. *The Quarterly Journal of Experimental Psychology Section A, 42*(3), 495-512. https://doi.org/10.1080/14640749008401234

Engle, R. W. (2001). *What is working memory capacity?* In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), The nature of remembering: Essays in honor of Robert G. Crowder (pp. 297–314). American Psychological Association. https://doi.org/10.1037/10394-016

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. *Models of working memory: Mechanisms of active maintenance and executive control, 4*, 102-134. https://doi.org/10.1017/CBO9781139174909.007

Erdfelder, E., Auer, T. S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology, 217*(3), 108-124. https://doi.org/10.1027/0044-3409.217.3.108.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior research methods, instruments, & computers, 28*(1), 1-11. https://doi.org/10.3758/BF03203630

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M., & Walter, H. (2003). Emotional context modulates subsequent memory effect. *Neuroimage, 18*(2), 439-447. https://doi.org/10.1016/S1053-8119(02)00015-0

Eysenck, M. W. (1976). Arousal, learning, and memory. *Psychological bulletin, 83*(3), 389. https://doi.org/10.1037/0033-2909.83.3.389

Eysenck, M.W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences, 50*(7), 955-960. https://doi.org/10.1016/j.paid.2010.08.019

Eysenck, M.W., Derakshan, N., Santos, R., & Calvo, M.G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336. https://psycnet.apa.org/doi/10.1037/1528-3542.7.2.336

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic bulletin & review, 1*(4), 429-438. https://doi.org/10.3758/BF03210947

Fanselow, M. S., & Lester, L. S. (1988). *A functional behavioristic approach to aversively motivated behavior: predatory imminence as a determinant of the topography of defensive behavior*. In R. C. Bolles & M. D. Beecher (Eds.), Evolution and learning (pp. 185–212). Lawrence Erlbaum Associates, Inc.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is" special" about face perception?. *Psychological review, 105*(3), 482. https://doi.org/10.1037/0033-295X.105.3.482

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*, 175-191. https://doi.org/10.3758/BF03193146

Fendt, M., & Fanselow, M. S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neuroscience & Biobehavioral Reviews, 23*(5), 743-760. https://doi.org/10.1016/S0149-7634(99)00016-0

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General, 144*(1), 30. https://doi.org/10.1037/a0038211

Friedman, D., & Johnson Jr, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy research and technique, 51*(1), 6-28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Funayama, E. S., Grillon, C., Davis, M., & Phelps, E. A. (2001). A double dissociation in the affective modulation of startle in humans: effects of unilateral temporal lobectomy. *Journal of cognitive neuroscience, 13*(6), 721-729. https://doi.org/10.1162/08989290152541395

Galletly, C., Clark, C.R., McFarlane, A.C., & Weber, D.L. (2001). Working memory in posttraumatic stress disorder—an event-related potential study. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies, 14(*2), 295-309. https://doi.org/10.1023/A:1011112917797

Gardiner, J. M., & Java, R. I. (1993). Recognition memory and awareness: An experiential approach. *European Journal of Cognitive Psychology, 5*(3), 337-346. https://doi.org/10.1080/09541449308520122

Giesinger, J. M., Loth, F. L., Aaronson, N. K., Arraras, J. I., Caocci, G., Efficace, F., ... & EORTC Quality of Life Group. (2020). Thresholds for clinical importance were established to improve interpretation of the EORTC QLQ-C30 in clinical practice and research. *Journal of clinical epidemiology, 118*, 1-8. https://doi.org/10.1016/j.jclinepi.2019.10.003

Gilbert, P. (2001). Evolutionary approaches to psychopathology: The role of natural defences. *Australian & New Zealand Journal of Psychiatry, 35*(1), 17-27. https://doi.org/10.1046/j.1440-1614.2001.00856.x

Godden, D. R., & Baddeley, A. D. (1975). Context‐dependent memory in two natural environments: On land and underwater. *British Journal of psychology, 66*(3), 325-331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Golkar, A., Castro, V., & Olsson, A. (2015). Social learning of fear and safety is determined by the demonstrator's racial group. *Biology letters, 11*(1), 20140817. https://doi.org/10.1098/rsbl.2014.0817

Golkar, A., & Olsson, A. (2017). The interplay of social group biases in social threat learning. *Scientific reports, 7*(1), 1-5. https://doi.org/10.1038/s41598-017-07522-z

Goodman, J.B., Freeman, E.E., & Chalmers, K.A. (2019). The relationship between early life stress and working memory in adulthood: A systematic review and meta-analysis. *Memory, 27*(6), 868-880. https://doi.org/10.1080/09658211.2018.1561897

Grillon, C., & Charney, D. R. (2011). In the face of fear: Anxiety sensitizes defensive responses to fearful faces: Fearful faces are not feared. *Psychophysiology, 48*(12), 1745–1752. https://doi.org/10.1111/j.1469-8986.2011.01268.x

Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-Potentiated Startle in Humans: Effects of Anticipatory Anxiety on the Acoustic Blink Reflex. *Psychophysiology, 28*(5), 588–595. https://doi.org/10.1111/j.1469-8986.1991.tb01999.x

Grillon, C., & Davis, M. (1995). Acoustic startle and anticipatory anxiety in humans: Effects of monaural right and left ear stimulation. *Psychophysiology, 32*(2), 155-161. https://doi.org/10.1111/j.1469-8986.1995.tb03307.x

Gruber, O. (2001). Effects of domain-specific interference on brain activation associated with verbal working memory task performance. *Cerebral Cortex, 11*(11), 1047-1055. https://doi.org/10.1093/cercor/11.11.1047

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience, 14*(7), 488-501. https://doi.org/10.1038/nrn3524

Guinosso, S. A., Johnson, S. B., & Riley, A. W. (2016). Multiple adverse experiences and child cognitive development. *Pediatric research, 79*(1), 220-226. https://doi.org/10.1038/pr.2015.195

Gutchess, A. H., Ieuji, Y., & Federmeier, K. D. (2007). Event-related potentials reveal age differences in the encoding and recognition of scenes. *Journal of Cognitive Neuroscience, 19*(7), 1089-1103. https://doi.org/10.1162/jocn.2007.19.7.1089

Haaker, J., Golkar, A., Selbing, I., & Olsson, A. (2017). Assessment of social transmission of threats in humans using observational fear conditioning. *Nature Protocols, 12*(7), 1378-1386. https://doi.org/10.1038/nprot.2017.027

Hajcak, G., Klawohn, J., & Meyer, A. (2019). The utility of event-related potentials in clinical psychology. *Annual review of clinical psychology, 15*, 71-95. https://doi.org/10.1146/annurev-clinpsy-050718-095457

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in cognitive sciences, 5*(9), 394-400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hamm, A. O. (2020). Fear, anxiety, and their disorders from the perspective of psychophysiology. *Psychophysiology, 57*(2), e13474. https://doi.org/10.1111/psyp.13474

Hamm, A. O., Schupp, H. T., & Weike, A. I. (2003). *Motivational organization of emotions: Autonomic changes, cortical responses, and reflex modulation*. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), Handbook of affective sciences (pp. 187–211). Oxford University Press.

Hauffa, R., Rief, W., Brähler, E., Martin, A., Mewes, R., & Glaesmer, H. (2011). Lifetime traumatic experiences and posttraumatic stress disorder in the German population: results of a representative population survey. *The Journal of nervous and mental disease, 199*(12), 934-939. https://doi.org/ 10.1097/NMD.0b013e3182392c0d

Hautus, M.J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers, 27*(1), 46-51. https://doi.org/10.3758/BF03203619

Hautzinger, M., Keller, F., & Kühner, C. (2006). *BDI-II Beck Depressions-Inventar* [BDI-II Beck depression inventory]. Harcourt Test Services.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences, 4*(6), 223-233. https://doi.org/10.1016/S1364-6613(00)01482-0

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological psychiatry, 51*(1), 59-67. https://doi.org/10.1016/S0006-3223(01)01330-0

Heck, D.W., Arnold, N.R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods, 50*. 264-284. https://doi.org/10.3758/s13428-017-0869-7

Heim, C., & Nemeroff, C. B. (2001). The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biological psychiatry, 49*(12), 1023-1039. https://doi.org/10.1016/S0006-3223(01)01157-X

Hennings, A. C., Lewis-Peacock, J. A., & Dunsmoor, J. E. (2021). Emotional learning retroactively enhances item memory but distorts source attribution. *Learning & Memory, 28*(6), 178-186. https://doi.org/ 10.1101/lm.053371.120

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *Journal of Neuroscience, 19*(10), 3962-3972. https://doi.org/10.1523/JNEUROSCI.19-10-03962.1999

Herzog, J.I., & Schmahl, C. (2018). Adverse childhood experiences and the consequences on neurobiological, psychosocial, and somatic conditions across the lifespan. *Frontiers in psychiatry, 9*, 420. https://doi.org/10.3389/fpsyt.2018.00420

Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology, 126*(2), 193. https://doi.org/10.1037/a0025180

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience & Biobehavioral Reviews, 55*, 498-509. https://doi.org/10.1016/j.neubiorev.2015.06.002

Hoffman, A. N., Trott, J. M., Makridis, A., & Fanselow, M. S. (2022). Anxiety, fear, panic: An approach to assessing the defensive behavior system across the predatory imminence continuum. *Learning & Behavior,* 1-10. https://doi.org/10.3758/s13420-021-00509-x

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1140. https://doi.org/10.1037/a0015019

Iffland, B., Klein, F., Schindler, S., Kley, H., & Neuner, F. (2021). “She finds you abhorrent”- The impact of emotional context information on the cortical processing of neutral faces in depression. *Cognitive, Affective, & Behavioral Neuroscience, 21*(2), 426-444. https://doi.org/10.3758/s13415-021-00877-x

Ikkai, A., McCollough, A.W., & Vogel, E.K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of neurophysiology, 103*(4), 1963-1968. https://doi.org/10.1152/jn.00978.2009

Jacobs, W. J., & Nadel, L. (1998). Neurobiology of reconstructed memory. *Psychology, Public Policy, and Law, 4*(4), 1110. https://doi.org/10.1037/1076-8971.4.4.1110

Jaeger, A., Johnson, J. D., Corona, M., & Rugg, M. D. (2009). ERP correlates of the incidental retrieval of emotional information: effects of study-test delay. *Brain Research, 1269,* 105-113. https://doi.org/10.1016/j.brainres.2009.02.082

JASP Team (2022). JASP (Version 0.16.1)[Computer software].

Javidi, H., & Yadollahie, M. (2012). Post-traumatic stress disorder. *The international journal of occupational and environmental medicine, 3*(1), 2-9.

Jaworek, A., Weymar, M., Löw, A., & Hamm, A. O. (2014). Brain potentials reflecting spontaneous retrieval of emotional long-term memories. *Cognitive Neuroscience, 5*(3-4), 168-176. https://doi.org/10.1080/17588928.2014.916259

Jelinek, L., Moritz, S., Randjbar, S., Sommerfeldt, D., Püschel, K., & Seifert, D. (2008). Does the evocation of traumatic memories confound subsequent working memory performance in posttraumatic stress disorder (PTSD)? *Depression and Anxiety, 25*(2), 175-179. https://doi.org/10.1002/da.20300

Johnsen, G. E., & Asbjørnsen, A. E. (2008). Consistent impaired verbal memory in PTSD: A meta-analysis. *Journal of Affective Disorders 111*(1), 74-82. https://doi.org/10.1016/j.jad.2008.02.007

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3.

Johnson, M. K., Kounios, J., & Nolde, S. F. (1997). Electrophysiological brain activity and memory source monitoring. *NeuroReport, 8*(5), 1317-1320. https://doi.org/10.1097/00001756-199611250-00025

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory, 17*(5), 577-596. https://doi.org/10.1080/09658210902976969

Jovanovic, T., Kazama, A., Bachevalier, J., & Davis, M. (2012). Impaired safety signal learning may be a biomarker of PTSD. *Neuropharmacology, 62*(2), 695-704. https://doi.org/10.1016/j.neuropharm.2011.02.023

Jud, A., Rassenhofer, M., Witt, A., Münzer, A., & Fegert, J. M. (2017). *Häufigkeitsangaben zum sexuellen Missbrauch. Internationale Einordnung, Bewertung der Kenntnislage in Deutschland, Beschreibung des Entwicklungsbedarfs.* Berlin: Unabhängiger Beauftragter für Fragen sexuellen Kindesmissbrauchs.

Junghöfer, M., Elbert, T., Tucker, D. M., & Rockstroh, B. (2000). Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology, 37*(4), 523-532. https://doi.org/10.1111/1469-8986.3740523

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *Journal of Neuroscience, 24*(17), 4172-4180. https://doi.org/10.1523/JNEUROSCI.0624-04.2004

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience, 17*(11), 4302-4311.

Karl, A., Malta, L.S., & Maercker, A. (2006). Meta-analytic review of event-related potential studies in post-traumatic stress disorder. *Biological psychology, 71*(2), 123-147. https://doi.org/10.1016/j.biopsycho.2005.03.004

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association, 90*(430), 773-795.

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia, 39*(12), 1263-1276. https://doi.org/10.1016/S0028-3932(01)00116-6

Kausche, F. M., & Schwabe, L. (2020). Blocking under stress: Sustained attention to stimuli without predictive value?. *Neurobiology of Learning and Memory,* 107158. https://doi.org/10.1016/j.nlm.2020.107158

Keefe, R. S. E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research, 57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1.

Keightley, M. L., Chiew, K. S., Anderson, J. A., & Grady, C. L. (2010). Neural correlates of recognition memory for emotional faces and scenes. *Social Cognitive and Affective Neuroscience, 6*(1), 24-37. https://doi.org/10.1093/scan/nsq003

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words?. *Memory & cognition, 31*(8), 1169-1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language, 54*(1), 99-112. https://doi.org/10.1016/j.jml.2005.05.005

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science, 16*(4), 213-218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kensinger, E.A. (2009). Remembering the details: Effects of emotion. *Emotion review, 1*(2), 99-113. https://doi.org/10.1177%2F1754073908100432

Kensinger, E. A., Piguet, O., Krendl, A. C., & Corkin, S. (2005). Memory for contextual details: effects of emotion and aging. *Psychology and aging, 20*(2), 241.

Kissler, J., & Strehlow, J. (2017). Something always sticks? How emotional language modulates neural processes involved in face encoding and recognition memory. *Poznan Studies in Contemporary Linguistics, 53*(1), 63-93. https://doi.org/10.1515/psicl-2017-0004

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika, 75*(1), 70-98. https://doi.org/10.1007/s11336-009-9141-0

Klein, S. B. (2015). What memory is. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(1), 1-38. https://doi.org/10.1002/wcs.1333

Klein, F., Iffland, B., Schindler, S., Wabnitz, P., & Neuner, F. (2015). This person is saying bad things about you: the influence of physically and socially threatening context information on the processing of inherently neutral faces. *Cognitive, Affective, & Behavioral Neuroscience, 15*(4), 736-748. https://doi.org/10.3758/s13415-015-0361-8

Klumb, P. L. (1995). Cognitive failures and performance differences: validation studies of a German version of the cognitive failures questionnaire. *Ergonomics, 38*(7), 1456-1467. https://doi.org/10.1080/00140139508925202

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*(3), 557-577. https://doi.org/10.1017/S0048577201990559

Koso, M., & Hansen, S. (2006). Executive function and memory in posttraumatic stress disorder: a study of Bosnian war veterans. *European Psychiatry 21*(3), 167-173. https://doi.org/10.1016/j.eurpsy.2005.06.004

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences, 4*(12), 463-470. https://doi.org/10.1016/S1364-6613(00)01560-6

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54-64. https://doi.org/10.1038/nrn1825

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science, 9*(6), 490-493. https://doi.org/10.1111/1467-9280.00090

LaGarde, G., Doyon, J., & Brunet, A. (2010). Memory and executive dysfunctions associated with acute posttraumatic stress disorder. *Psychiatry research, 177*(1), 144-149. https://doi.org/10.1016/j.psychres.2009.02.002

Laing, P. A., & Harrison, B. J. (2021). Safety learning and the Pavlovian conditioned inhibition of fear in humans: Current state and future directions. *Neuroscience & Biobehavioral Reviews*, *127*, 659-674. https://doi.org/10.1016/j.neubiorev.2021.05.014

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior, 32*(1), 4-14. https://doi.org/10.3758/BF03196002

Lang, P. J., & McTeague, L. M. (2009). The anxiety disorder spectrum: Fear imagery, physiological reactivity, and differential diagnosis. *Anxiety, Stress, & Coping*, *22*(1), 5-25. https://doi.org/10.1080/10615800802478247

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: animal models and human cognitive psychophysiology. *Journal of affective disorders*, *61*(3), 137-159. https://doi.org/10.1016/S0165-0327(00)00343-8

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377-1388. https://doi.org/10.1080/02699930903485076

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State-Trait-Angstinventar* [The state-trait anxiety inventory]. Beltz Test GmbH.

Lavric, A., Rippon, G., & Gray, J. R. (2003). Threat-evoked anxiety disrupts spatial working memory performance: An attentional account. *Cognitive therapy and research, 27*(5), 489-504. https://doi.org/10.1023/A:1026300619569

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience, 23*(1), 155-184.

Lee, T. H., Sakaki, M., Cheng, R., Velasco, R., & Mather, M. (2014). Emotional arousal amplifies the effects of biased competition in the brain. *Social cognitive and affective neuroscience, 9*(12), 2067-2077. https://doi.org/10.1093/scan/nsu015

Letkiewicz, A.M., Silton, R.L., Mimnaugh, K.J., Miller, G.A., Heller, W., Fisher, J., & Sass, S.M. (2020). Childhood abuse history and attention bias in adults. *Psychophysiology, 57*(10), e13627. https://doi.org/10.1111/psyp.13627

Lindström, B. R., & Bohlin, G. (2012). Threat-relevance impairs executive functions: negative impact on working memory and response inhibition. *Emotion, 12*(2), 384. https://doi.org/10.1037/a0027305

Lindström, B., Haaker, J., & Olsson, A. (2018). A common neural network differentially mediates direct and social fear learning. *NeuroImage, 167*, 121-129. https://doi.org/10.1016/j.neuroimage.2017.11.039

Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., & Grillon, C. (2008). Generalization of conditioned fear-potentiated startle in humans: Experimental validation and clinical relevance. *Behaviour Research and Therapy, 46*(5), 678–687. https://doi.org/10.1016/j.brat.2008.02.005

Lobo, I., Portugal, L. C., Figueira, I., Volchan, E., David, I., Garcia Pereira, M., & de Oliveira, L. (2015). EEG correlates of the severity of posttraumatic stress symptoms: A systematic review of the dimensional PTSD literature. *Journal of Affective Disorders, 183*, 210–220. https://doi.org/10.1016/j.jad.2015.05.015

Luck, C. C., & Lipp, O. V. (2016). Instructed extinction in human fear conditioning: History, recent developments, and future directions. *Australian Journal of Psychology, 68*(3), 209-227. https://doi.org/10.1111/ajpy.12135

Luck, S. J., & Kappenman, E. S. (2012). *ERP components and selective attention*. In S. J. Luck & E. S. Kappenman (Eds.), The Oxford handbook of event-related potential components (pp. 295–327). Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281. https://doi.org/10.1038/36846

Luck, S.J., Vogel, E.K., & Shapiro, K.L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature, 383*(6601), 616-618. https://doi.org/10.1038/383616a0

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences, 4*(11), 432-440. https://doi.org/10.1016/S1364-6613(00)01545-X

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *Karolinska directed emotional faces* [Database of standardized facial images]. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 91, 630.

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and cognition, 65*(3), 209-237. https://doi.org/10.1016/j.bandc.2007.02.007

MacNamara, A., Ferri, J., & Hajcak, G. (2011). Working memory load reduces the late positive potential and this effect is attenuated with increasing anxiety. *Cognitive, Affective, & Behavioral Neuroscience, 11*, 321-331. https://doi.org/10.3758/s13415-011-0036-z

Mangels, J. A., Picton, T. W., & Craik, F. I. (2001). Attention and successful episodic encoding: an event-related potential study. *Cognitive Brain Research, 11*(1), 77-95. https://doi.org/10.1016/S0926-6410(00)00066-5

Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual review of neuroscience, 24*(1), 897-931. https://doi.org/10.1146/annurev.neuro.24.1.897

Martínez-Galindo, J. G., & Cansino, S. (2017). Emotional context during encoding modulates recognition electrophysiological activity. *Experimental Brain Research, 235*(1), 169-179. https://doi.org/10.1007/s00221-016-4780-8

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science, 2*(1), 33-52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language, 58*(2), 449-464. https://doi.org/10.1016/j.jml.2007.01.004

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on psychological science, 6*(2), 114-133. https://doi.org/10.1177/1745691611400234

Mathews, A. (1990). Why worry? The cognitive function of anxiety. *Behaviour research and therapy, 28*(6), 455-468. https://doi.org/10.1016/0005-7967(90)90132-3

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7.

Mattarozzi, K., Todorov, A., & Codispoti, M. (2015). Memory for faces: the effect of facial appearance and the context in which the face is encountered. *Psychological Research, 79*(2), 308-317. https://doi.org/10.1007/s00426-014-0554-8

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E. J. (2015). Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*, *80*(1), 205–235. https://doi.org/10.1007/s11336-013-9374-9.

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and aging, 20*(4), 571. https://doi.org/10.1037/0882-7974.20.4.571

McEwen, B.S. (2002). The neurobiology and neuroendocrinology of stress: Implications for post-traumatic stress disorder from a basic science perspective. *Psychiatric Clinics, 25*(2), 469-494. https://doi.org/10.1016/S0193-953X(01)00009-0

McGaugh, J.L. (2000). Memory--a century of consolidation. *Science, 287*(5451), 248-251. https://doi.org/10.1126/science.287.5451.248

McGrath, J. E. (1976). *Stress and behavior in organizations*. Handbook of industrial and organizational psychology, 1351, 1396.

McNally, R. J., Foa, E. B., & Donnell, C. D. (1989). Memory bias for anxiety information in patients with panic disorder. *Cognition & Emotion, 3*(1), 27-44. https://doi.org/10.1080/02699938908415234

McTeague, L. M., & Lang, P. J. (2012). The anxiety spectrum and the reflex physiology of defense: from circumscribed fear to broad distress. *Depression and anxiety, 29*(4), 264-281. https://doi.org/10.1002/da.21891

McTeague, L. M., Lang, P. J., Laplante, M. C., Cuthbert, B. N., Shumen, J. R., & Bradley, M. M. (2010). Aversive imagery in posttraumatic stress disorder: trauma recurrence, comorbidity, and physiological reactivity. *Biological Psychiatry*, *67*(4), 346-356. https://doi.org/10.1016/j.biopsych.2009.08.023

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. Clinical EEG and Neuroscience, 37(4), 292-299. https://doi.org/10.1177/155005940603700406

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & cognition, 34*(4), 865-876. https://doi.org/10.3758/BF03193433

Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(*1), 116. https://doi.org/10.1037/0278-7393.28.1.116

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(10), 980-989. https://doi.org/10.1016/j.jaac.2010.05.017

Mertens, G., Boddez, Y., Sevenster, D., Engelhard, I. M., & De Houwer, J. (2018). A review on the effects of verbal instructions in human fear conditioning: Empirical findings, theoretical considerations, and future directions. *Biological Psychology, 137*, 49-64. https://doi.org/10.1016/j.biopsycho.2018.07.002

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

Mikels, J. A., & Reuter-Lorenz, P. A. (2019). Affective working memory: An integrative psychological construct. *Perspectives on Psychological Science*, *14(*4), 543-559. https://doi.org/10.1177/1745691619837597

Mikels, J. A., Reuter-Lorenz, P. A., Beyer, J. A., & Fredrickson, B. L. (2008). Emotion and working memory: evidence for domain-specific processes for affective maintenance. *Emotion, 8*(2), 256. https://doi.org/10.1037/1528-3542.8.2.256

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience, 24*(1), 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory?. *Psychological bulletin, 135*(4), 638. https://doi.org/10.1037/a0015849

Mitte, K. (2008). Memory bias for threatening information in anxiety and anxiety disorders: a meta-analytic review. *Psychological bulletin, 134*(6), 886. https://doi.org/10.1037/a0013343

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive therapy and research, 29*(1), 29-45. https://doi.org/10.1007/s10608-005-1646-y

Moran, T.P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological bulletin, 142*(8), 831. https://psycnet.apa.org/doi/10.1037/bul0000051

Morgan, H.M., Klein, C., Boehm, S.G., Shapiro, K.L., & Linden, D.E. (2008). Working memory load for faces modulates P300, N170, and N250r. *Journal of cognitive neuroscience, 20*(6), 989-1002. https://doi.org/10.1162/jocn.2008.20072

Moriya, J., & Sugiura, Y. (2012). High visual working memory capacity in trait social anxiety. *PloS one, 7*(4), e34244. https://doi.org/10.1371/journal.pone.0034244

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: the ICE theory. *Journal of Experimental Psychology: General, 128*(4), 403. https://doi.org/10.1037/0096-3445.128.4.403

Näätänen, R. (1995). The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear and Hearing, 16*(1), 6-18.

Nairne, J. S., Pandeirada, J. N., & Thompson, S. R. (2008). Adaptive memory: The comparative value of survival processing. *Psychological Science, 19*(2), 176-180. https://doi.org/10.1111/j.1467-9280.2008.02064.x.

Natu, V., & O’Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British journal of psychology, 102*(4), 726-747. https://doi.org/10.1111/j.2044-8295.2011.02053.x

Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia, 50*(14), 3370-3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

Öğmen, H., & Herzog, M. H. (2016). A new conceptualization of human visual sensory-memory. *Frontiers in psychology, 7*, 830. https://doi.org/10.3389/fpsyg.2016.00830

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of personality and social psychology, 80*(3), 381. https://doi.org/10.1037/0022-3514.80.3.381

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford university press.

Olsson, A., McMahon, K., Papenberg, G., Zaki, J., Bolger, N., & Ochsner, K. N. (2016). Vicarious fear learning depends on empathic appraisals and trait empathy. *Psychological science, 27*(1), 25-33. https://doi.org/10.1177/0956797615604124

Olsson, A., Nearing, K. I., & Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. *Social cognitive and affective neuroscience, 2*(1), 3-11. https://doi.org/10.1093/scan/nsm005

Olsson, A., & Phelps, E. A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychological science, 15*(12), 822-828. https://doi.org/10.1111/j.0956-7976.2004.00762.x

Olsson, A., & Phelps, E.A. (2007). Social learning of fear. *Nature neuroscience, 10*(9), 1095-1102. https://doi.org/10.1038/nn1968

Omer, Y., Sapir, R., Hatuka, Y., & Yovel, G. (2019). What is a face? Critical features for face detection. *Perception, 48*(5), 437-446. https://doi.org/10.1177/0301006619838734

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*(1), 75-92. https://doi.org/10.1016/j.neuropsychologia.2006.04.025

Paret, C., & Bublatzky, F. (2020). Threat rapidly disrupts reward reversal learning. *Behaviour Research and Therapy, 103636*. https://doi.org/10.1016/j.brat.2020.103636

Pavlov, I.P. *Conditioned Reflexes* (Oxford Univ. Press, Oxford, 1927).

Pavlov, P. I. (2010). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of neurosciences, 17*(3), 136.

Pechtel, P., & Pizzagalli, D.A. (2011). Effects of early life stress on cognitive and affective function: an integrated review of human literature. *Psychopharmacology, 214*(1), 55-70. https://doi.org/10.1007/s00213-010-2009-2

Pelosi, L., Slade, T., Blumhardt, L. D., & Sharma, V. K. (2000). Working memory dysfunction in major depression: an event-related potential study. *Clinical neurophysiology, 111*(9), 1531-1543. https://doi.org/10.1016/S1388-2457(00)00354-0

Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: A meta-analysis. *Clinical psychology review, 35*, 10-18. https://doi.org/10.1016/j.cpr.2014.10.005

Pessoa, L., & Ungerleider, L. G. (2004). Neural correlates of change detection and change blindness in a working memory task. *Cerebral Cortex, 14*(5), 511-520. https://doi.org/10.1093/cercor/bhh013

Peyk, P., DeCesarei, A., & Junghöfer, M. (2011). Electro Magneto Encephalograhy Software: overview and integration with other EEG/MEG toolboxes. *Computational Intelligence and Neuroscience, 2011*. https://doi.org/10.1155/2011/861705

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature neuroscience, 4*(4), 437-441. https://doi.org/10.1038/86110

Piccardi, L., Nori, R., Boccia, M., Barbetti, S., Verde, P., Guariglia, C., & Ferlazzo, F. (2015). A dedicated system for topographical working memory: Evidence from domain-specific interference tests. *Experimental Brain Research, 233*(8), 2489-2495. https://doi.org/10.1007/s00221-015-4320-y

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience, 13*(1), 25-42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual review of neuroscience, 35,* 73-89. https://doi.org/10.1146/annurev-neuro-062111-150525

Potts, G. F., & Tucker, D. M. (2001). Frontal evaluation and posterior representation in target detection. *Cognitive Brain Research, 11*(1), 147-156. https://doi.org/10.1016/S0926-6410(00)00075-6

Quaedflieg, C. W., & Schwabe, L. (2018). Memory dynamics under stress. *Memory, 26*(3), 364-376. https://doi.org/10.1080/09658211.2017.1338299

Qureshi, S. U., Long, M. E., Bradshaw, M. R., Pyne, J. M., Magruder, K. M., Kimbrell, T., ... & Kunik, M. E. (2011). Does PTSD impair cognition beyond the effect of trauma?. *The Journal of neuropsychiatry and clinical neurosciences, 23*(1), 16-28.

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour research and therapy, 15*(5), 375-387. https://doi.org/10.1016/0005-7967(77)90041-9

Raes, A. K., De Houwer, J., De Schryver, M., Brass, M., & Kalisch, R. (2014). Do CS-US pairings actually matter? A within-subject comparison of instructed fear conditioning with and without actual CS-US pairings. *PloS one, 9*(1), e84888. https://doi.org/10.1371/journal.pone.0084888

Rak, N., Bellebaum, C., & Thoma, P. (2013). Empathy and feedback processing in active and observational learning. *Cognitive, Affective, & Behavioral Neuroscience, 13*(4), 869-884. https://doi.org/10.3758/s13415-013-0187-1

Ramon, M., Caharel, S., & Rossion, B. (2011). The speed of recognition of personally familiar faces. *Perception, 40*(4), 437-449. https://doi.org/10.1068/p6794

Read, T. R. C., & Cressie, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data*. New York, NY: Springer.

Robinson, O. J., Charney, D. R., Overstreet, C., Vytal, K., & Grillon, C. (2012). The adaptive threat bias in anxiety: Amygdala–dorsomedial prefrontal cortex coupling and aversive amplification. *NeuroImage, 60*(1), 523–529. https://doi.org/10.1016/j.neuroimage.2011.11.096

Robinson, O.J., Vytal, K., Cornwell, B.R., & Grillon, C. (2013). The impact of anxiety upon cognition: perspectives from human threat of shock studies. *Frontiers in human neuroscience, 7*, 203. https://doi.org/10.3389/fnhum.2013.00203

Roozendaal, B., McEwen, B. S., & Chattarji, S. (2009). Stress, memory and the amygdala. *Nature Reviews Neuroscience, 10*(6), 423-433. https://doi.org/10.1038/nrn2651

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic bulletin & review, 12*(4), 573-604. https://doi.org/10.3758/BF03196750

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of mathematical psychology*, *56*(5), 356-374. https://doi.org/10.1016/j.jmp.2012.08.001

Rowles, M. E., Lipp, O. V., & Mallan, K. M. (2012). On the resistance to extinction of fear conditioned to angry faces. *Psychophysiology, 49*(3), 375-380. https://doi.org/10.1111/j.1469-8986.2011.01308.x

Ruchkin, D. S., Johnson Jr, R., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: An event-related potential study. *Cognitive Brain Research*, *1*(1), 53-66. https://doi.org/10.1016/0926-6410(92)90005-C

Rugg, M.D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004.

Saar-Ashkenazy, R., Shalev, H., Kanthak, M.K., Guez, J., Friedman, A., & Cohen, J.E. (2015). Altered processing of visual emotional stimuli in posttraumatic stress disorder: an event-related potential study. *Psychiatry Research: Neuroimaging, 233*(2), 165-174. https://doi.org/10.1016/j.pscychresns.2015.05.015

Sauro, M. D., Jorgensen, R. S., & Teal Pedlow, C. (2003). Stress, glucocorticoids, and memory: a meta-analytic review. *Stress, 6*(4), 235-245. https://doi.org/10.1080/10253890310001616482

Savitz, J., van der Merwe, L., Stein, D. J., Solms, M., & Ramesar, R. (2007). Genotype and childhood sexual trauma moderate neurocognitive performance: a possible role for brain-derived neurotrophic factor and apolipoprotein E variants. *Biological psychiatry, 62*(5), 391-399. https://doi.org/10.1016/j.biopsych.2006.10.017

Schaper, M. L., Kuhlmann, B. G., & Bayen, U. J. (2019). Metamemory expectancy illusion and schema-consistent guessing in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(3), 470–496. https://doi.org/10.1037/xlm0000602.

Schellhaas, S., Arnold, N., Schmahl, C., & Bublatzky, F. (2020). Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study. *Neurobiology of Learning and Memory, 174,* 107280. https://doi.org/10.1016/j.nlm.2020.107280

Schellhaas, S., Schmahl, C, & Bublatzky, F. (under review [a]). Social learning in individuals with adverse childhood experiences: Memory processes as a function of instructional and observational threat and safety learning [Manuscript submitted for publication].

Schellhaas, S., Schmahl, C., & Bublatzky, F. (under review [b]). Incidental learning of faces during threat: No evidence for increased autonomic arousal to “unrecognized” threat identities [Manuscript submitted for publication].

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex*. https://doi.org/10.1016/j.cortex.2020.06.010

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology, 33*(5), 643-653. https://doi.org/10.1016/j.psyneuen.2008.02.004

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in brain research, 156*, 31-51. https://doi.org/10.1016/S0079-6123(06)56002-9

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *NeuroReport, 14*(8), 1107-1110. https://doi.org/10.1097/00001756-200306110-00002

Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: an ERP analysis. *Emotion, 4(*2), 189-200. https://doi.org/10.1037/1528-3542.4.2.189

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2008). The selective processing of emotional visual stimuli while detecting auditory targets: An ERP analysis. *Brain Research, 1230*, 168-176. https://doi.org/10.1016/j.brainres.2008.07.024

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *Journal of Neuroscience, 27*(5), 1082-1089. https://doi.org/10.1523/JNEUROSCI.3223-06.2007

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: an update and integration. *Neuroscience & Biobehavioral Reviews, 36*(7), 1740-1749. https://doi.org/10.1016/j.neubiorev.2011.07.002

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex, 80*, 141-153. https://doi.org/10.1016/j.cortex.2015.11.001

Selbing, I., Lindström, B., & Olsson, A. (2014). Demonstrator skill modulates observational aversive learning. *Cognition, 133*(1), 128-139. https://doi.org/10.1016/j.cognition.2014.06.010

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 1005-1025. https://doi.org/10.1037/0278-7393.24.4.1005

Sessa, P., Luria, R., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2011). Interhemispheric ERP asymmetries over inferior parietal cortex reveal differential visual working memory maintenance for fearful versus neutral facial identities. *Psychophysiology, 48*(2), 187-197. https://doi.org/10.1111/j.1469-8986.2010.01046.x

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric, A., & Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion, 6*(1), 40. https://doi.org/10.1037/1528-3542.6.1.40

Shackman, A. J., Stockbridge, M. D., Tillman, R. M., Kaplan, C. M., Tromp, D. P. M., Fox, A. S., & Gamer, M. (2016). The Neurobiology of Dispositional Negativity and Attentional Biases to Threat: Implications for Understanding Anxiety Disorders in Adults and Youth. *Journal of Experimental Psychopathology, 7*(3), 311–342. https://doi.org/10.5127/jep.054015

Shea, A., Walsh, C., MacMillan, H., & Steiner, M. (2005). Child maltreatment and HPA axis dysregulation: relationship to major depressive disorder and post traumatic stress disorder in females. *Psychoneuroendocrinology, 30*(2), 162-178. https://doi.org/10.1016/j.psyneuen.2004.07.001

Sheridan, M.A., & McLaughlin, K.A. (2014). Dimensions of early experience and neural development: deprivation and threat. *Trends in cognitive sciences, 18*(11), 580-585. https://doi.org/10.1016/j.tics.2014.09.001

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in cognitive sciences, 9*(1), 16-20. https://doi.org/10.1016/j.tics.2004.11.006

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE, 94*(11), 1948-1962. https://doi.org/ 10.1109/JPROC.2006.884093

Smith, A. P., Dolan, R. J., & Rugg, M. D. (2004). Event-related potential correlates of the retrieval of emotional and nonemotional context. *Journal of Cognitive Neuroscience, 16*(5), 760-775. https://doi.org/10.1162/089892904970816

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology, 54*(1), 167-183. https://doi.org/10.1016/j.jmp.2009.06.007

Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Frontiers in psychology, 6*, 328. https://doi.org/10.3389/fpsyg.2015.00328

Solomon, S. D., & Davidson, J. R. (1997). Trauma: prevalence, impairment, service use, and cost. *The Journal of clinical psychiatry, 58*, 5-11.

Spataro, J., Mullen, P. E., Burgess, P. M., Wells, D. L., & Moss, S. A. (2004). Impact of child sexual abuse on mental health. *The British Journal of Psychiatry, 184*(5), 416-421. https://doi.org/10.1192/bjp.184.5.416

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor perspectives in biology, 7*(8), a021766. https://doi.org/ 10.1101/cshperspect.a021766

Stangier, U., & Steffens, M. (2001). *Social Phobia Inventory—deutsche Fassung* [German version]. Psychologisches Institut der Universität Frankfurt am Main.

Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., & Hoyer, J. (1999). Die Erfassung sozialer Phobie durch Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS) [Assessment of social phobia by the Social Interaction Anxiety Scale (SIAS) and the Social Phobia Scale (SPS)]. *Zeitschrift für klinische Psychologie, 28*(1), 28–36. https://doi.org/10.1026//0084-5345.28.1.28

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers, 31*(1), 137-149. https://doi.org/10.3758/BF03207704

Stegmann, Y., Ahrens, L., Pauli, P., Keil, A., & Wieser, M.J. (2020). Social aversive generalization learning sharpens the tuning of visuocortical neurons to facial identity cues. *Elife, 9*, e55204. https://doi.org/10.7554/eLife.55204.sa2

Stein, M. B., Kennedy, C. M., & Twamley, E. W. (2002). Neuropsychological function in female victims of intimate partner violence with and without posttraumatic stress disorder. *Biological psychiatry, 52*(11), 1079-1088. https://doi.org/10.1016/S0006-3223(02)01414-2

Stout, D.M., Shackman, A.J., & Larson, C.L. (2013). Failure to filter: anxious individuals show inefficient gating of threat from working memory. *Frontiers in human neuroscience, 7,* 58. https://doi.org/10.3389/fnhum.2013.00058

Sussman, T. J., Jin, J., & Mohanty, A. (2016). Top-down and bottom-up factors in threat-related perception and attention in anxiety. *Biological psychology, 121*, 160-172. https://doi.org/10.1016/j.biopsycho.2016.08.006

Sutherland, M. R., & Mather, M. (2012). Negative arousal amplifies the effects of saliency in short-term memory. *Emotion, 12*(6), 1367. https://doi.org/10.1037/a0027860

Symeonidou, N., & Kuhlmann, B. G. (2022). Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory. *Cognition and Emotion, 36*(2), 300-316. https://doi.org/10.1080/02699931.2021.2008323

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science, 22*(6), 430-436. https://doi.org/10.1177/0963721413498893

Tapia, G., Clarys, D., Bugaiska, A., & El-Hage, W. (2012). Recollection of negative information in posttraumatic stress disorder. *Journal of traumatic stress, 25*(1), 120-123. https://doi.org/10.1002/jts.21659

Tepas, D. I., & Price, J. M. (2000). *What is stress and what is fatigue*?. In Stress, workload, and fatigue (pp. 607-622). CRC Press.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1016. https://doi.org/10.1037/0096-1523.25.4.1016

Tottenham, N., Tanaka, J.W., Leon, A.C., McCarry, T., Nurse, M., Hare, T.A., ... & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research, 168*(3), 242-249. https://doi.org/10.1016/j.psychres.2008.05.006

Tulving, E. (1983). Ecphoric processes in episodic memory. *Philosophical Transactions of the Royal Society of London. B 302*(1110), 361-371. https://doi.org/10.1098/rstb.1983.0060

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological review, 80*(5), 352. https://doi.org/10.1037/h0020071

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language, 67*(1), 1-16. https://doi.org/10.1016/j.jml.2011.12.005

van Ast, V. A., Spicer, J., Smith, E. E., Schmer-Galunder, S., Liberzon, I., Abelson, J. L., & Wager, T. D. (2014). Brain mechanisms of social threat effects on working memory. *Cerebral Cortex, 26*(2), 544-556. https://doi.org/10.1093/cercor/bhu206

van Rooij, S. J., & Jovanovic, T. (2019). Impaired inhibition as an intermediate phenotype for PTSD risk and treatment response. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 89*, 435-445. https://doi.org/10.1016/j.pnpbp.2018.10.014

Vasterling, J. J., Brailey, K., Constans, J. I., & Sutker, P. B. (1998). Attention and memory dysfunction in posttraumatic stress disorder. *Neuropsychology, 12*(1), 125. https://doi.org/10.1037/0894-4105.12.1.125

Vasterling, J. J., Duke, L. M., Brailey, K., Constans, J. I., Allain Jr, A. N., & Sutker, P. B. (2002). Attention, learning, and memory performances and intellectual resources in Vietnam veterans: PTSD and no disorder comparisons. *Neuropsychology, 16*(1), 5. https://doi.org/ 10.1037//0894-4105.16.1.5

Ventura-Bort, C., Löw, A., Wendt, J., Dolcos, F., Hamm, A. O., & Weymar, M. (2016a). When neutral turns significant: brain dynamics of rapidly formed associations between neutral stimuli and emotional contexts. *European Journal of Neuroscience, 44*(5), 2176-2183. https://doi.org/10.1111/ejn.13319

Ventura-Bort, C., Löw, A., Wendt, J., Moltó, J., Poy, R., Dolcos, F., ... & Weymar, M. (2016b). Binding neutral information to emotional contexts: brain dynamics of long-term recognition memory. *Cognitive, Affective, & Behavioral Neuroscience, 16*(2), 234-247. 10.3758/s13415-015-0385-0

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500-503. https://doi.org/10.1038/nature04171

Vormbrock, F., & Neuser, J. (1983). Konstruktion zweier spezifischer Trait-Fragebogen zur Erfassung von Angst in sozialen Situationen (SANB und SVSS). *Diagnostica, 29*(2), 165-182.

Vytal, K., Cornwell, B., Arkin, N., & Grillon, C. (2012). Describing the interplay between anxiety and cognition: from impaired performance under low cognitive load to reduced anxiety under high load. *Psychophysiology, 49*(6), 842-852. https://doi.org/10.1111/j.1469-8986.2012.01358.x

Wald, I., Degnan, K. A., Gorodetsky, E., Charney, D. S., Fox, N. A., Fruchter, E., ... & Bar-Haim, Y. (2013). Attention to threats and combat-related posttraumatic stress symptoms: prospective associations and moderation by the serotonin transporter gene. *JAMA psychiatry, 70*(4), 401-408. https://doi.org/ 10.1001/2013.jamapsychiatry.188

Wallace, J. C., Kass, S. J., & Stanny, C. J. (2002). The cognitive failures questionnaire revisited: dimensions and correlates. *The Journal of general psychology, 129*(3), 238-256. https://doi.org/10.1080/00221300209602098

Wangelin, B. C., Bradley, M. M., Kastner, A., & Lang, P. J. (2012). Affective engagement for facial expressions and emotional scenes: The influence of social anxiety. *Biological Psychology, 91*(1), 103–110. https://doi.org/10.1016/j.biopsycho.2012.05.002

Ward, R.T., Lotfi, S., Sallmann, H., Lee, H.J., & Larson, C.L. (2020). State anxiety reduces working memory capacity but does not impact filtering cost for neutral distracters. *Psychophysiology, 57*(10), e13625. https://doi.org/10.1111/psyp.13625

Weymar, M., & Hamm, A. O. (2013). *Electrophysiological signature of emotional memories*. In M. Linden, & K. Rutkowski (Eds.), Hurting Memories and Beneficial Forgetting: Postraumatic Stress Disorders, Biographical Developments and Social Conflicts (pp. 21-35). Elsevier.

Weymar, M., Bradley, M. M., Hamm, A. O., & Lang, P. J. (2013). When fear forms memories: Threat of shock and brain potentials during encoding and recognition. *Cortex, 49*(3), 819-826. https://doi.org/10.1016/j.cortex.2012.02.012

Weymar, M., Bradley, M. M., Hamm, A. O., & Lang, P. J. (2014). Encoding and reinstatement of threat: recognition potentials. *Neurobiology of Learning and Memory, 107*, 87-92. https://doi.org/10.1016/j.nlm.2013.11.005

Weymar, M., Löw, A., & Hamm, A. O. (2011). Emotional memories are resilient to time: evidence from the parietal ERP old/new effect. *Human Brain Mapping, 32*(4), 632-640. https://doi.org/10.1002/hbm.21051

Wieser, M. J., & Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in psychology, 3*, 471. https://doi.org/10.3389/fpsyg.2012.00471

Wieser, M. J., Gerdes, A. B., Büngel, I., Schwarz, K. A., Mühlberger, A., & Pauli, P. (2014). Not so harmless anymore: How context impacts the perception and electrocortical processing of neutral faces. *NeuroImage, 92*, 74-82. https://doi.org/10.1016/j.neuroimage.2014.01.022

Wieser, M. J., Pauli, P., Reicherts, P., & Mühlberger, A. (2010). Don't look at me in anger! Enhanced processing of angry faces in anticipation of public speaking. *Psychophysiology, 47*(2), 271-280. https://doi.org/10.1111/j.1469-8986.2009.00938.x

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia, 37*(4), 441-454. https://doi.org/10.1016/S0028-3932(98)00100-6

Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology, 35*(1), 81-87. https://doi.org/10.1016/S0167-8760(99)00095-1

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain, 119*(3), 889-905. https://doi.org/10.1093/brain/119.3.889

Wilding, E. L., & Rugg, M. D. (1997). An event-related potential study of memory for words spoken aloud or heard. *Neuropsychologia, 35*(9), 1185-1195. https://doi.org/10.1016/S0028-3932(97)00048-1

Wilding, E. L., Doyle, M. C., & Rugg, M. D. (1995). Recognition memory with and without retrieval of context: An event-related potential study. *Neuropsychologia, 33*(6), 743-767. https://doi.org/10.1016/0028-3932(95)00017-W

Wilkinson, D., & Halligan, P. (2004). The relevance of behavioural measures for functional-imaging studies of cognition. *Nature Reviews Neuroscience, 5*(1), 67-73. https://doi.org/10.1038/nrn1302

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders*. Second Edition. Chichester: Wiley.

Wolf, O. T. (2009). Stress and memory in humans: twelve years of progress?. *Brain research, 1293*, 142-154. https://doi.org/10.1016/j.brainres.2009.04.013

Young, A. W., & Burton, A. M. (2017). Recognizing Faces. Current Directions in *Psychological Science, 26*(3), 212–217. https://doi.org/10.1177/0963721416688114

Publications

Peer-reviewed

Bublatzky, F., Schellhaas, S., & Guerra, P. (2022). The mere sight of loved ones does not inhibit psychophysiological defense mechanisms when threatened. Scientific reports, 12(1), 1-10. https://doi.org/10.1038/s41598-022-06514-y

Arnold, N. R., González Cruz, H., Schellhaas, S., & Bublatzky, F. (2021). A multinomial modelling approach to face identity recognition during instructed threat. Cognition and Emotion, 35(7), 1302-1319. https://doi.org/10.1080/02699931.2021.1951175

Schellhaas, S., Arnold, N., Schmahl, C., & Bublatzky, F. (2020). Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study. Neurobiology of Learning and Memory, 174, 107280. https://doi.org/10.1016/j.nlm.2020.107280

Vieira, J. B., Schellhaas, S., Enström, E., & Olsson, A. (2020). Help or flight? Increased threat imminence promotes defensive helping in humans. Proceedings of the Royal Society B, 287(1933), 20201473. https://doi.org/10.1098/rspb.2020.1473

Schredl, M., Berres, S., Klingauf, A., Schellhaas, S., & Göritz, A. S. (2015). Factors affecting the frequency of music dreams: An online study. International Journal of Dream Research, 8(2), 139-141.

Schredl, M., Berres, S., Klingauf, A., Schellhaas, S., & Göritz, A. S. (2014). The Mannheim Dream Questionnaire (MADRE): Retest reliability, age and gender effects. International Journal of Dream Research, 7(2), 141-147.

Under review

Schellhaas, S., Schmahl, C, & Bublatzky, F. (under review). Social learning in individuals with adverse childhood experiences: Memory processes as a function of instructional and observational threat and safety learning [European Journal of Psychotraumatology].

Bublatzky, F., Schellhaas, S., & Paret, C. (under review). Comparing apples and oranges: Electrocortical correlates of decision-making and reward reversal learning under threat [Frontiers in Behavioral Neuroscience].

Schellhaas, S., Schmahl, C., & Bublatzky, F. (under review). Incidental learning of faces during threat: No evidence for increased autonomic arousal to “unrecognized” threat identities [Scientific Reports].

Conference contributions (Posters and Talks)

Schellhaas, S., Schmahl, C., & Bublatzky, F. (2022, May). Sensing danger: Autonomic arousal to recognized and unrecognized person identity from formerly threatening or safe contexts. Poster session presented at the Social & Affective Neuroscience Society (SANS) virtual conference.

Schellhaas, S., Schmahl, C., & Bublatzky, F. (2022, March). Incidental learning of person identity: Autonomic arousal to “unrecognized” threat identities. In N. Symeonidou & H. Tanyas (Chairs), Source Memory. Symposium conducted at the 64th Conference of Experimental Psychologists (TeaP), Cologne, Germany (online conference)

Schellhaas, S., Schmahl, C., & Bublatzky, F. (2021, October). Face perception and recognition as a function of threat, adverse childhood experiences and anxiety. In I. Niedtfeld (Chair), Investigating social cognition in individuals with different levels of personality pathology and childhood trauma. Symposium conducted at the conference of the International Society for the Study of Personality Disorders (ISSPD), Oslo, Norway (online conference).

Schellhaas, S., Schmahl, C., & Bublatzky, F. (2021, June). The impact of adverse childhood experiences on social learning, face perception and recognition: An ERP study. In F. Bublatzky & S. Schindler (Chairs), Person perception as a function of attention, emotion and learning history: Recent findings from electrophysiology. Symposium conducted at the conference of the European Society for Cognitive and Affective Neuroscience (ESCAN), Budapest, Hungary (online conference).

Schellhaas, S., Schmahl, C., & Bublatzky, F. (2021, May). The interplay of adverse childhood experiences and social threat learning on face perception and recognition: ERP studies. In U. Lüken & J. Richter (Chairs), Clinical psychology meets neuroscience: basic mechanisms, endophenotypes and transdiagnostic perspectives. Symposium conducted at the conference of the “Fachgruppe Klinische Psychologie und Psychotherapie der DGPS“, Mannheim, Germany (online conference)

Schellhaas, S., Arnold, N. R., Schmahl, C., & Bublatzky, F. (2020, October). The impact of adverse childhood experiences and social threat learning on electrocortical face processing and recognition. Virtual poster session presented at the annual meeting of the Society for Psychophysiological Research (SPR).

Schellhaas, S., Arnold, N. R., Schmahl, C., & Bublatzky, F. (2019, September). Face identity learning during instructed threat: Electrocortical evidence from a source memory task. In Y. G. Pavlov & C. Panitz (Chairs), Brain and body in danger: central and peripheral electrophysiology of stimulus processing under threat. Symposium conducted at the annual meeting of the Society for Psychophysiological Research (SPR), Washington, DC, USA.

Schellhaas, S., Arnold, N. R., Schmahl, C., & Bublatzky, F. (2019, June). Encoding and recognition of faces as a function of contextual settings: An ERP study. In A. Schacht & M. Weymar (Chairs), Emotion and Motivation as Context: New Evidence for Biased Processing in Perception, Learning and Memory. Symposium conducted at the “Psychologie und Gehirn“ (Psychology and Brain) conference, Dresden, Germany.

Schellhaas, S., Paret, C., Schmahl, C., & Bublatzky, B. (2018, May/June). Error monitoring under threat: Electrocortical correlates of learning reward contingencies. Poster session presented at the “Psychologie und Gehirn“ (Psychology and Brain) conference, Gießen, Germany.

Schellhaas, S., Arnold, N. R., Schmahl, C., & Bublatzky, F. (2019, May). Incidental learning of face identity during threat: An ERP study. Poster session presented at the European Meeting of Human Fear Conditioning (EMHFC), Würzburg, Germany.

List Of Tables

List Of Figures

Curriculum Vitae

Sabine Schellhaas

born 7th August 1991 in Bad Friedrichshall, Germany

Department of Psychosomatic Medicine and Psychotherapy,
Central Institute of Mental Health Mannheim, University of Heidelberg

C4, 11, Room 406
68159 Mannheim (Germany)
(+49) (0) 621 1703 4461
sabine.schellhaas@zi-mannheim.de (sabine.schellhaas@googlemail.com)

Education

07/2018– present	**Doctoral (Dr. sc. hum) student** Zentralinstitut für Seelische Gesundheit Mannheim/Universität Heidelberg Research Group “Social Learning and Person Perception” Thesis topic: *The influence of contextual threat and safety transmitted by social learning on visual working memory and item/source memory and its modulation by adverse childhood experiences and (social) anxiety* Advisors: Prof. Dr. Christian Schmahl, Dr. Florian Bublatzky
09/2015– 12/2017	**Master of Science** (Cognitive and Clinical Psychology) University of Mannheim Thesis topic: *Electrocortcal correlates of (reversal) learning of reward and loss contingencies: A threat-of-shock study* Advisors: Prof. Dr. Georg W. Alpers, Dr. Florian Bublatzky
15/12/2017	Final grade: 1,2
09/2012– 09/2015	**Bachelor of Science** (Psychology) University of Mannheim Thesis topic: *Gender differences in self-perceived causes of test anxiety* Advisors: Prof. Dr. Oliver Dickhäuser, Dr. Ann Krispenz (née Seibert)
3/9/2015	Final grade: 1,7
2002 – 2011	**Eduard-Mörike-Gymnasium Neuenstadt am Kocher**
31/05/2011	Abitur: 1,2

Academic positions

02/2018– present	**Research assistant** in the research group "Social learning and person perception" (DFG funded), Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health Mannheim Advisors: Dr. Florian Bublatzky, Prof. Dr. Christian Schmahl
01/2020– 03/2020	**Research stay** Emotion Lab Karolinska Institutet Stockholm, Sweden Advisors: Prof. Dr. Andreas Olsson, Dr. Joana Vieira
10/2017– 01/2018	**Student research assistant** Department of Psychosomatic Medicine and Psychotherapy Central Institute of Mental Health Mannheim
05/2017– 07/2017	**Intern at the Psychiatric Ambulance for Cognitive Training (PAKT)** Department for General Psychiatry University Hospital Heidelberg Advisor: Claudia Bach, M.A.
03/2017– 05/2017	**Student research assistant** Chair for Clinical and Biological Psychology and Psychotherapy University of Mannheim
02/2014– 01/2017 & 08/2017– 09/2017	**Student research assistant** Chair for Cognitive Psychology and Individual Differences University of Mannheim

Honors and awards

01/2019– 12/2021	**Doctoral scholarship** granted by the Studienstiftung des deutschen Volkes
04/2018– 06/2021	**Affiliation with the graduate school GRK 2350** (DFG funded) Central Institute of Mental Health Mannheim Spokesperson: Prof. Dr. Christian Schmahl

Acknowledgements

First and foremost, I would like to thank my supervisors Dr. Florian Bublatzky and Prof. Dr. Christian Schmahl who sparked my interest in the social learning of threat, introduced me to the application of psychophysiological and neural measurement methods and broadened my scientific horizons in so many ways. I appreciate Christian's quick and insightful comments on my work and his goal-oriented problem-solving skills, which moved my projects forward. Christian, thank you for accepting and sharing my excitement for source memory.

This thesis would not have been possible without Florian's support and guidance, who saw something in me during my master's studies and encouraged me to pursue this academic path. I greatly appreciate his enthusiasm in research coupled with his calm and precise way of thinking and working, all of which has influenced my research immensely. Florian, thank you for always giving me the space to delevolp myself and my research independently, but also offering help when I needed it.

I would like to thank all the student research assistants who helped me with data collection and who provided me with company during the long hours in the lab. A special thank goes to Max, Tana, and Niclas. I would also like to thank all the individuals who participated in the studies of this thesis for their time, trust, and patience.

I was fortunate to be a part of the great research training group GRK2350. I value the exchange with the members, especially the doctoral students, and the resulting influence on my thesis projects very much. The association to the GRK gave me the opportunity to work with two great mentors, Prof. Dr. Katja Bertsch and Dr. Nathalie Holz. Their advice on career-related topics, but also on work-life balance and shared experiences on working in academia shaped my way of thinking of who I want to be as a (female) researcher. Marija, Angelika, Sylvia, and Julia have been a great support for the administrative organization of the studies and thesis and offered help beyond their area of responsibility. Marija, you were the best office companion during the tedious corona times that I could have asked for!

My time at (and outside of) the office would not have been half as enjoyable without Malte, Miroslava, Nike, Maurizio, and Sophia! Malte, thank you for your excellent cooking skills and your thoughtful gestures and words of encouragement in difficult times. Mirus, thank you for always being willing to listen to my complaints, and your great advice. Nike, thank you for your positivity, the great conversations and for having the courage to make a bet with me where the chances of you winning it were zero! Maurizio, thank you for being the best office mate, for your contagious passion for science, and for teaching me how to use the white board properly. Sophia, thank you for becoming such a valuable friend in such a short time, for your openness and directness and your support in finishing this thesis.

A special thanks goes to Alfred, who I could always bother with EEG-related problems. Thank you for your emotional support and for challenging me to be the best version of myself. And to my family and friends at home, Andreas, Florian, Katrin, Bastian, Felix, Stefanie, and Lisa, who have shaped me from the beginning, always encouraged me and helped me to look at life from different perspectives. Thank you for being there for me!

Friends are the family you choose – and this could not be more true than for my dear roommate of 7 years, Nikoletta. Thank you for always building me up, believing in me and for allowing me to be who I really am. I am beyond grateful to have you in my life!

Last but not least, I would like to thank my parents, Johanna and Heinrich Schellhaas, without whom none of this would have been possible. Their constant and unconditional support and excitement for everything in my life is limitless. Mama, Papa, es gibt kaum Worte, die beschreiben, was eure Unterstützung für mich bedeutet. Einfach danke.