JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

Neuronale Korrelate der Bedrohungsverarbeitung bei Sozialer Angststörung

Inauguraldissertation
zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)
vorgelegt von

Susanne Fricke, M. Sc. Psychologie

an der
Justus-Liebig-Universität Gießen
Fachbereich 06: Psychologie und Sportwissenschaften
Professur für Psychotherapie und Systemneurowissenschaften
Otto-Behaghel-Straße 10 H
35394 Gießen

Gießen, April 2026

Erstgutachterin: Apl. Prof. Dr. Andrea Hermann

Zweitgutachter: Prof. Dr. Benjamin Straube

Erklärung

„Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nichtveröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten sowie ethische, datenschutzrechtliche und tierschutzrechtliche Grundsätze befolgt. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt und indirekt an der Entstehung der vorliegenden Arbeit beteiligt waren. Mit der Überprüfung meiner Arbeit durch eine Plagiatserkennungssoftware bzw. ein internetbasiertes Softwareprogramm erkläre ich mich einverstanden.“

Gießen, den 01.04.2026 _____

Inhaltsverzeichnis

Abbildungsverzeichnis

Studie 1

Studie 2

Tabellenverzeichnis

Studie 1

Studie 2

Überblick

Die Soziale Angststörung (SAD) ist geprägt von verschiedenen Aspekten dysfunktionaler Bedrohungsverarbeitung, die sich beispielsweise in Prozessen des Lernens und Verlernens von Furcht (z. B. Ahrens et al., 2016; Savage et al., 2020), nachteiligen Aufmerksamkeitsprozessen und verzerrten Bewertungsmustern (Kim et al., 2018; Lidle & Schmitz, 2021) äußert. Für diese Dissertation wurden zwei Forschungsschwerpunkte identifiziert. Da SAD und die Posttraumatische Belastungsstörung (PTBS) ähnliche Symptomprofile sowie gemeinsame Defizite in Furchtkonditionierungsprozessen aufweisen, wurde eine Studie in Anlehnung an PTBS-Befunde (Garfinkel et al., 2014; Milad et al., 2009; Wicking et al., 2016) konzipiert. Dabei wurde die kontextabhängige Modulation konditionierter Furcht und deren neuronaler Korrelate mittels funktioneller Magnetresonanztomographie (fMRT) bei Patient:innen mit SAD untersucht (Fricke et al., 2024). Es wurde auch ein Vergleich von Patient:innen mit und ohne Intrusionssymptome angestellt. Ein zweiter Schwerpunkt stellte die selbstbezogene Verarbeitung als Aspekt der dysfunktionalen Bedrohungsverarbeitung dar. Dabei wurden neutrale Gesichter als soziale Reize genutzt, welche bei SAD häufig negativ interpretiert werden (Lange et al., 2012). Zunächst wird überblicksartig ein theoretischer Hintergrund zu den beiden Forschungsschwerpunkten mit dem bisherigen Forschungsstand gegeben. Anschließend werden die Fragestellungen und damit verbundenen Hypothesen dargestellt. Zu den beiden Fragestellungen sind jeweils die entsprechenden Artikel in Manuskriptform eingefügt. Im Anschluss findet sich eine weiterführende Diskussion der Befunde und deren Implikationen unter Berücksichtigung der Limitationen der vorliegenden Studien.

Theoretischer Hintergrund

Die Soziale Angststörung (SAD) ist definiert durch die Angst vor sozialen Situationen, in denen Individuen befürchten, negativ bewertet zu werden (American Psychiatric Association, 2022; H.-J. Yoon et al., 2019). Soziale Situationen werden in der Folge häufig vermieden oder unter großer Anspannung erduldet, sodass aus dieser Erkrankung weitreichende negative Folgen für das Funktionsniveau und die Lebensqualität der

Betroffenen resultieren (Aderka et al., 2012). Bei der Entstehung und Aufrechterhaltung der SAD spielen Prozesse dysfunktionaler Bedrohungsverarbeitung eine bedeutsame Rolle. Diese äußern sich unter anderem in einem Aufmerksamkeitsbias gegenüber Bedrohungen (Kim et al., 2018), einer stärkeren emotionalen Reaktivität und schlechteren kognitiven Regulation bei sozialer Bedrohung (Goldin, Manber et al., 2009) und negativer antizipatorischer, sowie post-ereignisbezogener Verarbeitung (Lidle & Schmitz, 2021). Auch Rapee und Heimberg (1997) berücksichtigen in ihrem Erklärungsmodell dysfunktionale Bedrohungsverarbeitung. Sie beschreiben, dass ein Aufmerksamkeitsfokus von Patient:innen mit SAD auf externen Reizen liegt, welche möglicherweise Anhaltspunkte für eine negative Bewertung durch das Gegenüber sein könnten (z. B. in Falten gezogene Stirn des Gegenübers als Zeichen dafür, dass er/sie gelangweilt ist).

Für die Entwicklung einer SAD spielen außerdem auslösende, sozial aversive Lernerfahrungen, sogenannte „soziale Traumata“, eine wichtige Rolle (Norton & Abbott, 2017; Wong & Rapee, 2016). Dies können beispielsweise das Bloßgestellt werden vor der Klasse, ein Kommentar über die roten Wangen oder ein Missgeschick bei einem Auftritt sein. In der Folge können sich die Symptome der SAD, aber auch andere Symptome, welche eine phänomenologische Ähnlichkeit zur Posttraumatischen Belastungsstörung (PTBS) aufweisen, entwickeln (Bjornsson et al., 2020; Erwin et al., 2006). Bjornsson et al. (2020) berichten in einer Stichprobe von 60 Patient:innen mit SAD von circa einem Drittel, welches die Kriterien für eine PTBS oder klinisch relevante posttraumatische Stresssymptome (Erfüllung einzelner Kriterien der PTBS) in Reaktion auf ein „soziales Trauma“ erfüllt. In Erklärungsmodellen der SAD werden negativ verzerrte Vorstellungsbilder von sich selbst beschrieben (z. B. sich selbst übergroß im Raum stehend; sich selbst mit knallroten Wangen, „wie eine Tomate“) (D. M. Clark & Wells, 1995; Heimberg et al., 2014). Diese enthalten meist Aspekte der auslösenden Situation und werden, wie Intrusionen auch, häufig ungewollt wiedererlebt und durch entsprechende soziale Situationen ausgelöst (Bjornsson et al., 2020; A. Hackmann et al., 2000; Ann Hackmann et al., 1998; Moscovitch et al., 2018).

Intrusionen, welche nach (sozial) aversiven Ereignissen auftreten, können als konditionierte Furchtreaktionen interpretiert werden (Franke et al., 2021). Für die Entwicklung und Aufrechterhaltung verschiedener Angststörungen und PTBS sind insbesondere veränderte Furchtkonditionierungsprozesse, welche ebenfalls einen Aspekt der dysfunktionalen Bedrohungsverarbeitung darstellen, relevant (Cooper et al., 2022; Duits et al., 2015; Lissek & Grillon, 2012; Lissek et al., 2005; Mineka & Zinbarg, 2006; Wong & Rapee, 2016). In der Literatur werden für SAD während der Furchtakquisition eine erhöhte Furchtreaktion – sowohl auf konditionierte Reize, als auch auf Sicherheitssignale – und Defizite während des Extinktionslernens berichtet (Ahrens et al., 2016; Chauret et al., 2019; C. Hermann et al., 2002; Lissek & Grillon, 2012; Lissek et al., 2008; Marin et al., 2020; Rabinak et al., 2017; Savage et al., 2020; Tinoco-González et al., 2015). Allerdings zeigen sich diese Befunde nicht konsistent über alle Studien und Ergebnismaße hinweg. Das neuronale Netzwerk der Furchtkonditionierung umfasst mehrere Schlüsselregionen. Während die Amygdala primär an der emotionalen Bewertung beteiligt ist, integriert der Hippocampus kontextuelle Informationen. Die regulatorische Kontrolle erfolgt modulierend durch den ventromedialen Präfrontalkortex (vmPFC) und exzitatorisch durch den dorsalen anterioren cingulären Kortex (dACC). Zusätzlich ist die Insula an der kortikalen Repräsentation der Furcht beteiligt (Bouton et al., 2006; A. Hermann et al., 2016; Kalisch et al., 2006; Maren et al., 2013; Motzkin et al., 2015; Phelps et al., 2001; Sehlmeyer et al., 2009). Studien zeigen, dass diese neuronalen Prozesse der Furchtkonditionierung bei Patient:innen mit Angststörungen und PTBS verändert sind (Chauret et al., 2019; Garfinkel et al., 2014; Marin et al., 2017; Milad et al., 2009; Suarez-Jimenez et al., 2020; VanElzakker et al., 2018).

Ein zentraler Faktor für das Wiederauftreten konditionierter Furcht bei PTBS ist die unzureichende Kontextualisierung der traumatischen Erfahrung, was dazu führt, dass Furchtreaktionen auch in sicheren Kontexten ausgelöst werden (Ehlers & Clark, 2000). In Laborstudien konnten entsprechende Defizite und physiologische/neuronale Veränderungen in kontextabhängigen Furchtkonditionierungsparadigmen festgestellt werden (Garfinkel et al., 2014; Milad et al., 2009; Wicking et al., 2016). Bei den Patient:innen mit PTBS zeigte sich das

Wiederauftreten der konditionierten Furcht im sicheren Extinktionskontext, ein geringeres Renewal im Gefahrenkontext und stärkeres Auftreten der Furcht in einem neuen Kontext, im Vergleich zu einer gesunden Kontrollgruppe. Die Übertragung von Erkenntnissen aus der PTBS-Forschung auf SAD erscheint vielversprechend, da beide Erkrankungen gemeinsame ätiologische und symptomatische Merkmale aufweisen. Insbesondere die Bedeutung auslösender traumatischer Ereignisse und posttraumatische Stresssymptome wie Intrusionen sprechen für transdiagnostische Mechanismen. Daher erscheint es plausibel, dass auch bei SAD Defizite bei der kontextabhängigen Furchtmodulation eine Rolle spielen könnten. Die Studienlage zu neuronalen Korrelaten kontextabhängiger Furchtkonditionierung ist für SAD sehr begrenzt. Es gibt eine fMRT Studie, welche während des Extinktionsabrufs in einem sicheren Kontext die reduzierte Aktivierung des vmPFC im Vergleich zu gesunden Kontrollprobanden zeigt, allerdings ohne Unterschiede bezüglich der Erregung (gemessen über die elektrodermale Reaktion) (Marin et al., 2020). Die Frage danach, ob auch bei SAD eine dysfunktionale Nutzung des Kontextes zum Wiederauftreten konditionierter Furcht führt, bleibt bisher unzureichend beantwortet.

Die dysfunktionale Bedrohungsverarbeitung bei SAD erstreckt sich nicht nur auf konditionierte, also eindeutig bedrohliche Reize, sondern auch auf uneindeutige soziale Reize. Ein Beispiel hierfür ist die negative Verarbeitung neutraler Gesichter, die insbesondere von sozial ängstlichen Menschen (Peschard & Philippot, 2017; K. L. Yoon & Zinbarg, 2008) und Patient:innen mit SAD (Lange et al., 2012) als bedrohlich wahrgenommen werden. Neutrale Gesichter werden bisher häufig als Kontrollbedingung bei der Untersuchung von Verarbeitungsprozessen emotionaler Gesichtsausdrücke genutzt (Gilboa-Schechtman & Shachar-Lavie, 2013; Machado-de-Sousa et al., 2010), gleichwohl sie demnach keine geeignete Vergleichsbedingung darstellen. Untersuchungen, welche neutrale Gesichter mit „verpixelten“ („scrambled“) Bildern von Gesichtern (nicht mehr als solche erkennbar), Fixationsbedingungen (z. B. Fixationsoval) oder einem aversiven Geruch verglichen, ergaben eine stärkere Aktivierungen der Amygdala und des fusiformen Gyrus bei Patient:innen mit SAD im Vergleich zu gesunden Kontrollen (Birbaumer et al., 1998; Cooney et al., 2006; Gentili et

al., 2008). Diese neuronalen Auffälligkeiten, speziell die stärkere Amygdala-Aktivierung könnten wiederum eine Erscheinungsform der verstärkten Bedrohungsverarbeitung bei SAD darstellen. Allerdings wurden auch bei diesen Untersuchungen innerhalb der Paradigmen emotionale Gesichter, bzw. aversive Reize (Geruch) präsentiert, sodass die Reaktion auf die neutralen Gesichter eventuell beeinflusst sein könnte. Interessant wäre daher die Untersuchung der Verarbeitung neutraler Gesichter bei SAD in einem Paradigma ohne konfundierende Stimuli, wie emotionale Gesichter.

Erklärungsmodelle der SAD beschreiben, dass Patient:innen solche (uneindeutigen) sozialen Reize, beziehungsweise allgemein externen Reize, als möglichen Hinweis für eine negative Bewertung der eigenen Person nutzen (Heimberg et al., 2014). Demnach könnte eine Tendenz zur selbstbezogenen Verarbeitung bei der negativen Interpretation neutraler Gesichter eine Rolle spielen. Selbstbezogene Verarbeitung beschreibt den kognitiven Prozess, bei dem Individuen externe Informationen im Hinblick auf ihre eigene Person evaluieren (z. B. „Tom ist so schnell an mir vorbei gelaufen - wahrscheinlich mag er mich nicht“) (Northoff, 2011; Northoff et al., 2011). Wenn Menschen ihre Aufmerksamkeit stark auf sich selbst richten, kann dies soziale Angst verstärken, denn es entstehen mehr negative Gedanken, positive äußere Reize werden seltener wahrgenommen und die eigene Leistung leidet darunter (Burgio et al., 1986; Norton & Abbott, 2016b; Spurr & Stopa, 2002; Woody & Rodriguez, 2000). Die neuronale Basis selbstbezogener Prozesse umfassen die sogenannten kortikalen Mittellinienstrukturen (orbitomedialer Präfrontalkortex, dorsomedialer Präfrontalkortex, anteriorer cingulärer Kortex [ACC], Precuneus/posteriorer cingulärer Kortex [PCC]; Northoff et al., 2006). Studien zur selbstbezogenen Verarbeitung bei SAD zeigen eine Hyperaktivität der kortikalen Mittellinienstrukturen, des temporalen Pols (TP), der temporo-parietalen Verbindung (TPJ) und der Insula (H.-J. Yoon et al., 2019). Dabei wurde der Selbstbezug beispielsweise über die Variation selbstbezogener oder auf andere bezogener Kommentare, das Lesen von Kommentaren aus der Ersten- oder Zweiten-Personen-Perspektive oder das Betrachten von Gesichtern mit oder ohne Augenkontakt manipuliert (Blair et al., 2008; Blair et al., 2011; Goldin & Gross, 2010; Goldin, Ramel & Gross, 2009;

Schneier et al., 2011; H.-J. Yoon et al., 2016). Unklar ist bisher, ob eine Erhöhung des Selbstbezugs die Wahrnehmung neutraler Gesichter bei Patient:innen mit SAD negativ beeinflusst und ob dies mit den bekannten neuronalen Veränderungen einhergeht.

Fragestellungen

Ziel dieser Dissertation war die Untersuchung verschiedener Aspekte der Bedrohungsverarbeitung und deren neuronaler Korrelate bei SAD.

Fragestellung Studie 1: Zeigen sich bei Patient:innen mit SAD ähnliche Defizite bei der kontextabhängigen Modulation von konditionierter Furcht, wie bei Patient:innen mit PTBS?

Ein Aspekt der Bedrohungsverarbeitung bei SAD zeigt sich in veränderten Furchtkonditionierungsprozessen, welche zur Entstehung und Aufrechterhaltung von sozialen Ängsten beitragen können (Chauret et al., 2019; C. Hermann et al., 2002; Lissek et al., 2008; Rabinak et al., 2017). Während Defizite in der kontextabhängigen Modulation von Furcht, insbesondere das Wiederauftreten von Furcht in sicheren und neuen Kontexten und deren neuronale Korrelate bei PTBS bereits mehrfach belegt wurden (Garfinkel et al., 2014; Milad et al., 2009; Wicking et al., 2016), wurde dies bei Patient:innen mit SAD noch nicht ausreichend untersucht. Eine erste Studie gibt allerdings Hinweise darauf, dass auch bei SAD eine dysfunktionale kontextabhängige Modulation konditionierter Furcht eine Rolle für die Aufrechterhaltung der Angst spielt (Marin et al., 2020). Auch phänomenologische Ähnlichkeiten zwischen SAD und PTBS, in Form eines ätiologisch relevanten auslösenden Ereignisses und posttraumatischen Stresssymptomen, wie intrusives Wiedererleben, legen das Vorliegen ähnlicher Defizite in der Kontextualisierung des ursprünglichen Ereignisses/ der Erinnerung daran bei SAD wie bei PTBS nahe. Um die kontextabhängige Modulation von Furcht zu untersuchen, wurde in der Dissertationsstudie mit Patient:innen mit SAD und gesunden Kontrollprobanden (HC) ein kontextabhängiges Furchtkonditionierungsparadigma mit fünf experimentellen Phasen durchgeführt (Akquisition in Kontext A, Extinktionstraining in Kontext B, Extinktionsabruf in Kontext B, Furchtrenewal in Kontext C und Furchtrenewal in Kontext A). Währenddessen wurde die neuronale Aktivität in furcht- und

extinktionsassoziierten Arealen und gleichzeitig die elektrodermale Reaktion erfasst. Es wurden Gruppenvergleiche zwischen SAD und HC sowie zwischen Patient:innen mit und ohne klinisch relevante Intrusionssymptomen durchgeführt.

Hypothesen Studie 1: 1. Im sicheren Extinktionskontext B – reduzierter Extinktionsabruf in der SAD-Gruppe, erkennbar an reduzierter differentieller konditionierter Aktivierung im Hippocampus und vmPFC, stärkerer Aktivierung im dACC, in der Insula und der Amygdala sowie stärkerer differentieller konditionierter elektrodermaler Aktivität. 2. Im neuen Kontext C und im Akquisitionskontext A – geringeres Furchtrenewal in der SAD-Gruppe, erkennbar an reduzierter differentieller konditionierter Aktivierung in der Amygdala, im dACC, in der Insula und im Hippocampus, stärkerer Aktivierung im vmPFC sowie reduzierter differentieller konditionierter elektrodermaler Reaktion. 3. Diese Defizite sollten sich vor allem bei den Patient:innen mit Intrusionen im Vergleich zu Patient:innen ohne diese Symptomatik zeigen. 4. Die Gruppen sollten sich auch bezüglich der Konnektivität in den untersuchten Hirnarealen unterscheiden.

Fragestellung Studie 2: Werden neutrale Gesichter von Patient:innen mit SAD im Vergleich zu HC aversiver wahrgenommen und führt eine Verstärkung des Selbstbezugs wiederum zu einer stärkeren negativen Reaktion bei Patient:innen mit SAD?

Eine defizitäre Bedrohungsverarbeitung zeigt sich bei Patient:innen mit SAD nicht nur bei eindeutig bedrohlichen Reizen, sondern beispielsweise auch bei uneindeutigen sozialen Reizen, wie neutralen Gesichtern (Lange et al., 2012). Obwohl die Verarbeitung von Gesichtern bei SAD bereits zuvor erforscht wurde, lag der Fokus hierbei meist auf emotionalen Gesichtern (Gilboa-Schechtman & Shachar-Lavie, 2013; Machado-de-Sousa et al., 2010). Die neuronalen Korrelate der Verarbeitung neutraler Gesichter sind bisher nicht hinreichend untersucht worden. Erklärungsmodelle der SAD legen nahe, dass Patient:innen mit SAD dazu tendieren, solche sozialen Hinweisreize auf sich zu beziehen (Heimberg et al., 2014). Demnach ist die selbstbezogene Verarbeitung möglicherweise ein Faktor, der zur negativen Wahrnehmung von uneindeutigen sozialen Hinweisreizen wie neutralen Gesichtern führt und

stellt gleichzeitig einen Aspekt der defizitären Bedrohungsverarbeitung bei SAD dar. Um die Verarbeitung neutraler Gesichter und ihre neuronalen Korrelate zu untersuchen, wurde mit den Patient:innen mit SAD und den HC aus Studie 1 ein weiteres neu entwickeltes Paradigma mit neutralen Gesichtern mittels fMRT durchgeführt. Als Vergleichsbedingung dienten Bilder, bei denen die Pixel der Gesichter so vermischt waren, dass diese nicht mehr als Gesichter erkennbar waren. Die Proband:innen sollten die Gesichter, beziehungsweise vermischten Bilder, „aufmerksam betrachten“. Um zusätzlich den Einfluss der selbstbezogenen Verarbeitung zu untersuchen, gab es zwei weitere Bedingungen, mit denen der Selbstbezug manipuliert wurde. Die Proband:innen wurden aufgefordert, sich beim Betrachten der Gesichter vorzustellen, dass die Person über sie nachdenkt (Selbstbezug erhöht) oder an etwas anderes denkt (Selbstbezug reduziert).

Hypothesen: 1. Das Betrachten neutraler Gesichter sollte bei Patient:innen mit SAD im Vergleich zu HC mit einer reduzierten Valenz einhergehen – dabei stärkere Aktivierung von Amygdala und Hirnarealen, die mit selbstbezogener Verarbeitung zusammenhängen (mPFC, ACC, PCC, Precuneus, TPJ, Insula). 2. Ein verstärkter Selbstbezug sollte bei SAD im Vergleich zu HC zu einer stärkeren Reduktion der Valenz führen – dabei wurde eine stärkere Reduktion der Aktivierung von mit Selbstbezug assoziierten Arealen erwartet. 3. Die Reduktion des Selbstbezugs sollte zu Unterschieden in der Valenz und der Aktivierung der untersuchten Hirnareale führen.

Studie 1: Neural correlates of context-dependent extinction recall in social anxiety disorder: relevance of intrusions in response to aversive social experiences [1]

Abstract

Background: There are phenomenological similarities between social anxiety disorder (SAD) and posttraumatic stress disorder, such as a provoking aversive event, posttraumatic stress symptoms (e.g. intrusions) in response to these events and deficient (context-dependent) fear conditioning processes. This study investigated the neural correlates of context-dependent extinction recall and fear renewal in SAD, specifically in patients with intrusions in response to an etiologically relevant aversive social event.

Methods: During functional magnetic resonance imaging a two-day context-dependent fear conditioning paradigm was conducted in 54 patients with SAD and 54 healthy controls. This included fear acquisition (context A) and extinction learning (context B) on one day, and extinction recall (context B) as well as fear renewal (contexts C and A) one day later. The main outcome measures were blood oxygen level-dependent responses in regions of interest and skin conductance responses.

Results: Patients with SAD showed reduced differential conditioned amygdala activation during extinction recall in the safe extinction context and during fear renewal in the acquisition context compared to healthy controls. Patients with clinically relevant intrusions moreover exhibited hypoactivation of the ventromedial prefrontal cortex during extinction learning, extinction recall, and fear renewal in a novel context, while amygdala activation more strongly decreased during

[1] Studie 1 wurde publiziert unter Fricke, S., Seinsche, R. J., Neudert, M. K., Schäfer, A., Zehtner, R. I., Stark, R. & Hermann, A. (2024). Neural correlates of context-dependent extinction recall in social anxiety disorder: relevance of intrusions in response to aversive social experiences. *Psychological medicine*, *54*(3), 548–557. https://doi.org/10.1017/S0033291723002179. Das Urheberrecht für diesen Artikel liegt bei Cambridge University Press & Assessment, hier mit Genehmigung wiedergegeben ('Reprinted with permission').

extinction learning and increased during fear renewal in the acquisition context compared with patients without intrusions.

Conclusions: Our study provides first evidence that intrusions in SAD are associated with similar deficits in context-dependent regulation of conditioned fear via the ventromedial prefrontal cortex as previously demonstrated in posttraumatic stress disorder.

Introduction

Aversive social experiences are reported from a majority of patients with social anxiety disorder (SAD) (Erwin, Heimberg, Marx, & Franklin, 2006; McCabe, Antony, Summerfeldt, Liss, & Swinson, 2003; Norton & Abbott, 2017) and are identified as a crucial factor in the etiology of the disorder (Mineka & Zinbarg, 2006; Norton & Abbott, 2017; Wong & Rapee, 2016). In response to these etiologically relevant aversive social events, some patients with SAD suffer from posttraumatic stress symptoms (PTSS), such as reexperiencing/intrusive memories (Bjornsson et al., 2020; Erwin et al., 2006), which are known in the context of Post-traumatic stress disorder (PTSD) after traumatic events. Intrusion symptoms after an aversive social experience, for example in the form of negatively distorted self-images, are considered in explanatory models of SAD (Clark & Wells, 1995; Rapee & Heimberg, 1997), leading to dysfunctional shifts of attention that contribute to the maintenance of the disorder. They also cause the affected person to re-experience features of the original aversive social situation, even in novel social situations (Hackmann, Clark, & McManus, 2000). Obviously, there are some phenomenological similarities between SAD and PTSD like a triggering aversive event and subsequent PTSS, such as intrusions.

Intrusions can be understood as conditioned fear responses (Franke et al., 2021). Altered fear conditioning processes are assumed to underlie the development and maintenance of PTSD and anxiety disorders (Duits et al., 2015; Dvir, Horovitz, Aderka, & Shechner, 2019; Lissek et al., 2005; Lissek & Grillon, 2012; Mineka & Zinbarg, 2006; Wong & Rapee, 2016). Laboratory fear conditioning studies in patients with SAD indicate increased fear responses towards the

conditioned stimulus and the safety stimulus during fear acquisition, as well as deficient extinction learning, although this is not shown to be consistent across studies or across outcome measures within studies (Ahrens et al., 2016; Chauret et al., 2019; C. Hermann et al., 2002; Lissek et al., 2008; Marin et al., 2020; Rabinak et al., 2017; Savage et al., 2020; Tinoco-González et al., 2015).

On the neural level, the following brain regions are in particular involved in the modulation of conditioned fear expression via the amygdala: hippocampus (especially context-dependent modulation), insula (cortical representation of fear), ventromedial prefrontal cortex (vmPFC; inhibitory function), and dorsal anterior cingulate cortex (dACC; excitatory function) (Bouton, Westbrook, Corcoran, & Maren, 2006; Hermann, Stark, Milad, & Merz, 2016; Kalisch et al., 2006; Maren, Phan, & Liberzon, 2013; Phelps et al., 2001; Sehlmeyer et al., 2009). These neural processes of fear conditioning are altered in PTSD (Garfinkel et al., 2014; Milad et al., 2009; Suarez-Jimenez et al., 2020; VanElzakker, Staples-Bradley, & Shin, 2018) and anxiety disorders (Chauret et al., 2019; Marin et al., 2017). In a previous study, social anxiety in healthy participants was related to greater differential conditioned activation of the amygdala and hippocampus during fear acquisition, but lower activation during extinction learning, as well as a tendency for reduced vmPFC responses during extinction recall (Pejic, Hermann, Vaitl, & Stark, 2013).

In PTSD, return of fear even in safe contexts is of high clinical relevance, probably resulting from inadequate contextualization of the trauma memory (Ehlers & Clark, 2000). In line with this, patients with PTSD compared with trauma-exposed controls showed a stronger return of fear and lower activation of the vmPFC and hippocampus, as well as hyperactivity of the dACC during extinction recall in the safe extinction context (Milad et al., 2009). In a further study, reduced vmPFC activation and higher SCRs towards the conditioned stimulus (CS+), and stronger differential activation in the insula (CS+ vs. CS-) during extinction recall, but lower activation in the amygdala and vmPFC and less SCRs towards the CS+ during renewal in the acquisition context were found in patients with PTSD compared to combat controls (Garfinkel et al., 2014). Moreover, during fear renewal in a novel context, patients with PTSD had higher differential conditioned

hippocampal and amygdala activity as well as higher conditioned SCRs compared to healthy controls (HC) (Wicking et al., 2016). These results overall point to deficits in context-dependent fear conditioning in patients with PTSD, which may arise due to deficits in using contextual information for the modulation of fear.

There is only one fMRI study investigating context-dependent fear conditioning in patients with SAD (Marin, Hammoud, Klumpp, Simon, & Milad, 2020). They found hypoactivation of the vmPFC during extinction recall in the safe extinction context in SAD compared with HC, but no significant differences in SCRs. However, there are no studies up to date investigating the neural correlates of conditioned fear renewal in a novel or the acquisition context, nor the association between PTSS, such as intrusions and altered fear conditioning processes in SAD.

Therefore, we conducted a two-day context-dependent fear conditioning paradigm with fear acquisition (context A) and extinction learning (context B) on the first day, and extinction recall (context B) and fear renewal in a novel (context C) as well as the acquisition context (context A) approximately 24 hours later during functional magnetic resonance imaging (fMRI) in 54 patients with SAD and 54 HC. We expected reduced extinction recall in the safe context (reduced differential conditioned activation in hippocampus and vmPFC, stronger differential conditioned activation in dACC, insula and amygdala and stronger differential conditioned SCRs) and lower fear renewal in the novel and the acquisition context (reduced differential conditioned activation in amygdala, dACC, insula and hippocampus, stronger activation in vmPFC and reduced differential SCRs) in SAD compared to HC. These changes should be particularly evident in patients with intrusion symptoms compared to patients without intrusion symptoms. Moreover, these group differences should not only be reflected in altered activation but also in altered connectivity of fear- and extinction-relevant brain areas.

Methods and Materials

Subjects

Recruitment was conducted via the mailing list of the local university, the associated outpatient clinic, posters, flyers, and announcements in newspapers and on websites. In patients, SAD had to be the primary mental disorder; several disorders (e.g., borderline personality disorder, psychosis, PTSD), current substance abuse or use of psychotropic drugs and current psychotherapy were exclusion criteria. HC were required not to have a current or past mental disorder (for details regarding exclusion criteria see Supplementary Table 3). In the course of the study, there was a dropout of 18 patients with SAD and 12 HC (cancellation without reason [6 SAD/2 HC], scheduling reasons [3 SAD/3 HC], discomfort in scanner [4 SAD/2 HC], technical difficulties [3 SAD/4 HC], strong head movements during scanning [2 SAD/1 HC]). This resulted in a final sample of 54 patients and 54 HC who were matched for age, sex, and number of years of education. Table 1 shows relevant demographic as well as diagnostic information for both groups. An a priori power analysis for the comparison of patients with HC using G*Power (version 3.1.9.2, Faul, Erdfelder, Lang, & Buchner, 2007), determined a group size of at least n=51 within each group, in order to be able to identify a group difference (t-test) with a mean effect size of d=.5 (power (1-β)=0.8, α=.05). The study protocol was approved by the local ethics committee of the Department of Psychology and Sport Science at the Justus Liebig University Giessen and was conducted in accordance with the Declaration of Helsinki. All participants gave written informed consent.

Table 1. *Demographic and clinical variables for patients with social anxiety disorder and healthy controls*

variable	SAD *n*=54	HC *n*=54	test for group differences
age, *M*(*SD*)	27.17(6.09)	27.15(7.69)	
range (years)	18-48	20-57	*t*=.01(106), *p*=.989
sex, female/male	36/18	34/20	$\chi^2(1)=.16$, *p*=.687
years of education, *M*(*SD*)	17.94(3.24)	17.56(3.15)	*t*=6.2(106), *p*=.538
symptoms of SAD, *M*(*SD*)			
LSAS	64.49(20.21)	10.87(7.58)	*t*=18.09(66.4), *p*<.001
SPIN	38.46(9.43)	7.20(5.13)	*t*=21.39(81.9), *p*<.001
intrusion symptoms in response to an aversive social event (CAPS-5), *M*(*SD*)	1.44(.27)	0.11(.06)	*t*=4.79(59), *p*<.001
BDI-II, *M*(*SD*)	11.83(7.56)	3.00(3.91)	*t*=7.58(80.5), *p*<.001
current comorbid diagnoses (without personality disorders), %			
none	51.9	100	
one	35.1	0	
two or more	13.0	0	
current comorbid personality disorder diagnoses, %			
none	72.2	100	
one	25.9	0	
two	1.9	0	
psychotherapeutic treatment in the past, %	35.2	0	
psychotropic drugs in the past, %	9.3	0	

Note. Table 1 shows descriptive variables separated for groups (SAD=social anxiety disorder, HC=healthy controls) and test statistics for group differences. *T*-test, degrees of freedom (*df*), χ^2-statistics, significance level (*p*), LSAS=Liebowitz Social Anxiety Scale, SPIN=Social Phobia Inventory, CAPS-5=Clinician-administered PTSD scale, BDI-II=Beck Depression Inventory II, diagnoses were assessed by using the DIPS (Diagnostic Interview for Mental Disorders) and the SKID-II (Structured clinical interview for DSM-IV Axis II: Personality Disorders), CAPS-5=Clinician-administered PTSD scale, sample size varied for HC: CAPS-5 *n*=53, BDI *n*=51 and for SAD: LSAS *n*=53.

To investigate the relevance of intrusion symptoms for context-dependent fear conditioning processes in SAD, we divided the patients into two groups by using the raw score of criterion B (intrusion symptoms) of the Clinician-administered PTSD scale (CAPS-5; Cwik & Woud, 2015). As an intrusion score >1 is defined as clinically relevant, patients with clinically relevant intrusions (INT; raw score>1, *n*=19) and patients without intrusions (NO-INT; raw score=0, *n*=29) were grouped together, respectively. We excluded patients with a raw score of 1 (*n*=7) for these group comparisons. Table 2 shows relevant demographic as well as diagnostic information for both groups separately as well as group differences between groups.

Table 2 *Sociodemographic and clinical variables separated for groups of patients with clinically relevant intrusions in response to the aversive social event and patients without intrusions*

variable	INT *n*=19	NO-INT *n*=29	test for group differences
age, *M*(*SD*)	25.63(4.69)	28.03(7.08)	*t*=-1.30(46), *p*=.2
range (years)	20-38	18-48	
sex, female/male	11/8	19/10	χ^2(1)=.29, *p*=.594
years of education, *M*(*SD*)	17.21(2.78)	18.16(3.43)	*t*=-1.01(46), *p*=.32
symptoms of SAD, *M*(*SD*)			
LSAS	59.78(4.23)	64.79(3.76)	*t*=-0.86(45), *p*=.393
SPIN	38.26(9.21)	37.72(10.06)	*t*=.19(46), *p*=.849
intrusion symptoms in response to an aversive social event (CAPS-5), *M*(*SD*)	3.79(1.47)	0(0)	*t*=11.20(18), *p*<.001
depressive symptoms			
BDI-II, *M*(*SD*)	14.47(8.40)	8.90(6.21)	*t*=2.65(46), *p*=.011
Number of diagnoses of mood disorder, without mania, M(SD)	.32(.48)	.17(.47)	*t*=1.03(46), *p*=.309
current comorbid diagnoses (without personality disorders), %			*t*=1.37(46), *p*=.177
none	47.4	58.6	
one	31.5	34.5	
two or more	21.1	6.9	
current comorbid personality disorder diagnoses, %			*t*=.84(46), *p*=.405
none	68.4	79.3	
one	31.6	20.7	
two	0	0	
psychotherapeutic treatment in the past, %	31.6	31.0	χ^2(2)=1.61, *p*=.446
psychotropic drugs in the past, %	15.8	6.9	χ^2(1)=.97, *p*=.324

Note. INT=patients with clinically relevant intrusions, NO-INT=patients without intrusions, *T*-test, degrees of freedom (*df*), χ^2-statistics, significance level (*p*), LSAS=Liebowitz Social Anxiety Scale, SPIN=Social Phobia Inventory, BDI-II=Beck Depression Inventory II, diagnoses were assessed by using the DIPS (Diagnostic Interview for Mental Disorders) and the SKID-II (Structured clinical interview for DSM-IV Axis II: Personality Disorders), CAPS-5=Clinician-administered PTSD scale, sample size varied for INT: LSAS *n*=18.

Interviews and Questionnaires

To confirm the inclusion or exclusion criteria of social anxiety disorder and other diagnoses, two trained psychologists administered the Diagnostic Interview for Mental Disorders (DIPS; Margraf, Cwik, Suppiger, & Schneider, 2017). The Liebowitz Social Anxiety Scale (LSAS; Consbruch, Stangier, & Heidenreich, 2016; Fresco et al., 2001; Liebowitz, 1987) was used as an interview to assess symptom severity and as an additional criterion for inclusion in the SAD group or exclusion in the HC group (total score≥30). We additionally used the Social Phobia Inventory (SPIN; Connor et al., 2000; Sosic, Gieler, & Stangier, 2008) to validate inclusion/exclusion decisions. The total score had to be ≥25 (SAD) or <25 (HC).

To select an autobiographical event, we used a list of aversive social events, which was adapted from Erwin, Heimberg, Marx & Franklin (Erwin et al., 2006). The subsequent exploration of PTSS was referred to an aversive social event most likely to be associated with the onset of the disorder, or the worst social event participants have ever experienced. Two psychologists assessed associated PTSS (intrusion symptoms, avoidance, alterations in cognitions and mood and alterations in arousal and reactivity) in the past month via the CAPS-5 (Cwik & Woud, 2015; Spies et al., 2020).

Procedure

This study refers to data acquired within a larger project with six appointments (see Supplement A). Data reported here were collected on three appointments: first appointment with diagnostic and other interviews and the third and fourth appointment with the fear conditioning paradigm. The fMRI measurements took place within 14 days of the diagnostic interviews and followed each other for approximately 24 hours. Participation was compensated with ten euros/hour or course credit.

Context-dependent fear conditioning paradigm

The two-day context-dependent fear conditioning paradigm was conducted in the MRI Scanner. The electrical stimulation, which served as UCS, was delivered via electrodes (Coulbourn Transcuaneous Aversive Finger Stimulator, E13-22, Holliston Massachussets) attached to the fingertips of the index and middle fingers of the dominant hand. The strength of the electrical stimulation was individually adjusted at the beginning of the first MRI session so that it was very uncomfortable for the individual, but not yet painful. The groups did not differ in the intensity of adjusted electrical stimulation (SAD: *M*=1.80 mA, *SD*=.68; HC: *M*=1.77 mA, *SD*=.69; *t*(105)=.19, *p*=.79; range in both groups=.6-4 mA; INT: *M*=1.90 mA, *SD*=.60, range=1.1-4 mA; NO-INT: *M*=1.76 mA, *SD*=.77; *t*(45)=.66, *p*=.51; range=.6-4 mA).

The paradigm is adapted from previous studies (Hermann et al., 2016; Hermann et al., 2020; Milad et al., 2007). Pictures of different rooms (e.g. conference room) served as contexts. A lamp was placed in each of them and its light colors (yellow, blue) were used as reinforced conditioned (CS+) and non-reinforced (CS-) stimuli. During the first fMRI measurement, fear acquisition took place in context A (8 trials per CS type; 5 trials with UCS presentation), followed by extinction training in context B (16 trials per CS type). Approximately 24 hours later, retrieval testing was performed in the safe extinction context B, in a novel context C, and in the acquisition context A (each 8 trials per CS type). Each trial began with the picture of a context with the lamp turned off (3s), which then lighted yellow or blue (6s), followed by electrical stimulation (UCS) in reinforced CS+ trials during fear acquisition. The whole paradigm and trial structure is described in detail in Supplement B (see also Supplementary Figure 4).

After the second scanning session, participants filled in a post hoc rating (detailed information in Supplement C) to assess their knowledge about the different contexts (context recognition), as well as the expected probability for the occurrence of the electrical stimulation (UCS expectancy) in the three different contexts. On 9-point Likert scales, participants were asked to indicate valence, arousal, and fear in each context (turned-off lamp), respectively.

Functional Magnetic Resonance Imaging

We used a 3 Tesla whole body scanner with a 64 channel standard head coil (Siemens Prisma, Erlangen, Germany). A T2*-weighted gradient echo-planar imaging sequence (EPI) with 42 slices (slice thickness=3 mm; 0.75 mm gap; descending slice order; TE=30 ms; TR=2.5 s; flip angle=81°; Field of View=220x220 mm²; matrix-size=74x74 pixel, parallel imaging: GRAPPA acceleration 2; oblique slices 30 degrees tilt from the AC-PC line) covering the entire brain was used to assess BOLD responses. Three persons (two patients with SAD, one HC) were excluded from fMRI analyses due to a framewise displacement of more than 0.5 mm in at least 20% of the volumes of one scanning session. We used Statistical Parametric Mapping (SPM12, r7219, Wellcome Department of Cognitive Neurology, London, UK; 2009), implemented in Matlab R2019a (Mathworks Inc., Sherborn, MA, USA) for data preprocessing and analysis. After preprocessing and first-level analyses (for details see Supplement D), the following contrasts were calculated on an individual level: differential conditioned responses (CS+ minus CS-) for acquisition, extinction learning (first 8 compared with/and last 8 trials) and early recall/renewal (first 4 trials), respectively. In order two analyze differences between SAD and HC (both groups n=54) we conducted two-sample t-tests for the experimental phases of the conditioning paradigm (acquisition, early extinction, late extinction, extinction learning, early extinction recall, early fear renewal in the novel context and early fear renewal in the acquisition context) in second-level analyses. The same analyses were done for comparing the two groups of patients with clinically relevant intrusions (INT, n=19) and without intrusions (NO-INT, n=29). For the analyses of main effects, one sample t-tests for each experimental phase were performed with all participants in one group (n=108). Region of interest (ROI) analyses were planned a priori for amygdala, insula, hippocampus, dACC and vmPFC. We used the small volume correction option of SPM12 and a significance threshold of $\alpha \leq .05$ on voxel level with family-wise error (FWE) correction.The reported coordinates (x, y, z) refer to the MNI space.

To investigate the connectivity between the resulting (significant) voxels and the other fear-relevant brain areas (amygdala, hippocampus, vmPFC, dACC, insula) in the corresponding experimental conditions, psychophysiological interaction (PPI) analyses were performed. For this purpose, the corresponding voxels were used as seed region (volume of interest; 6mm sphere around the peak voxel), respectively.

Skin Conductance Responses

We placed Ag/AgCl electrodes filled with isotonic (0.05M NaCl) electrolyte medium on the hypothenar surface of the non-dominant hand to measure SCRs. Data were preprocessed as well as analyzed using Ledalab 3.4.4 (available under www.ledalab.de), a toolbox integrated in MATLAB R2018B (The MathWorks, Natick, MA, USA). We used the ‘trough-to-peak’ analyses (Pineles, Orr, & Orr, 2009). Detailed information is given in Supplement E. To correct for violation of normal distribution, we log(µS+1) transformed all responses.

Due to classification as non-responders (less than two SCRs >0.02 µS in reaction to the UCS after CS+ during fear acquisition, *n*=6) or technical problems during measurement (*n*=13), we had to exclude 19 persons, resulting in a sample of 48 SAD (INT *n*=17, NO-INT *n*=27) and 44 HC subjects for the SCR analyses (which also were included in fMRI data analyses). Statistical analyses were performed using IBM SPSS 28 (IBM Corporation, Armonk, USA). We conducted nine mixed ANOVAs for each of the two group comparisons, each with the within-subject factor ‘stimulus’ (CS+/CS-) and the between-subject factor ‘group’ (HC and SAD or INT and NO-INT). Analogous to the fMRI analyses, we conducted five ANOVAs for acquisition, 'late' extinction, and for the 'early' phases of extinction recall in the extinction context, renewal in the novel context, and renewal in the acquisition context, with these two factors. To investigate the decrease during extinction training, and especially the return of the conditioned SCRs between phases, we additionally performed four more ANOVAs. For these ANOVAs on extinction training and return of fear, each involving two phases, we added the within-subjects factor 'phase' (first/second half of extinction training; last two trials of the previous phase [subsequently referred to as 'late'] and the

first four trials of the following phase [subsequently referred to as 'early']: late extinction training/early extinction recall in the extinction context, late extinction recall/early renewal in the novel context, late renewal in the novel context/early renewal in the acquisition context).

Results

Comparison of patients with SAD and HC

Patients did not differ from HC in age, sex, and years of education, but, as expected, in severity of social anxiety and intrusion symptoms (CAPS-5) regarding their worst aversive social experience (see Table 1).

Analyses of main effects for context-dependent fear conditioning across both groups (SAD and HC) indicated successful fear conditioning and extinction learning, but also spontaneous recovery in the extinction context on the next day. In the novel, as well as acquisition context, fear renewal was shown (see Supplement F).

Mixed ANOVAs regarding SCRs during all phases of fear conditioning (see Figure 1), revealed a significant interaction between group and phase for early extinction recall in the extinction context compared with late extinction training (Greenhouse-Geisser $F_{(1,90)}$=4.09, p=.046). In Figure 1, patients with SAD appear to have a greater increase in activation in early extinction recall compared to late extinction training, although separate ANOVAs for each phase showed no significant main effects of group (all p≥.121). Regarding the other experimental phases, all other ANOVAs revealed no significant main effects of (all p≥.121) or interactions with (all p≥.108) group.

On the neural level, HC compared with patients showed increased differential conditioned activation (CS+>CS-) in the right amygdala during early extinction recall in the extinction context (T_{max}=3.52, $p_{FWEcorr}$=.024, MNI: x=30, y=-2, z=-22) and during early renewal in the acquisition context (T_{max}=3.31, $p_{FWEcorr}$=.039, MNI: x=32, y=-2, z=-22) (see Figure 2). PPI analyses with these right amygdala activation foci as seed regions revealed no significant group differences in

connectivity with activation in any ROIs. There were no differences at the neural level in the other phases.

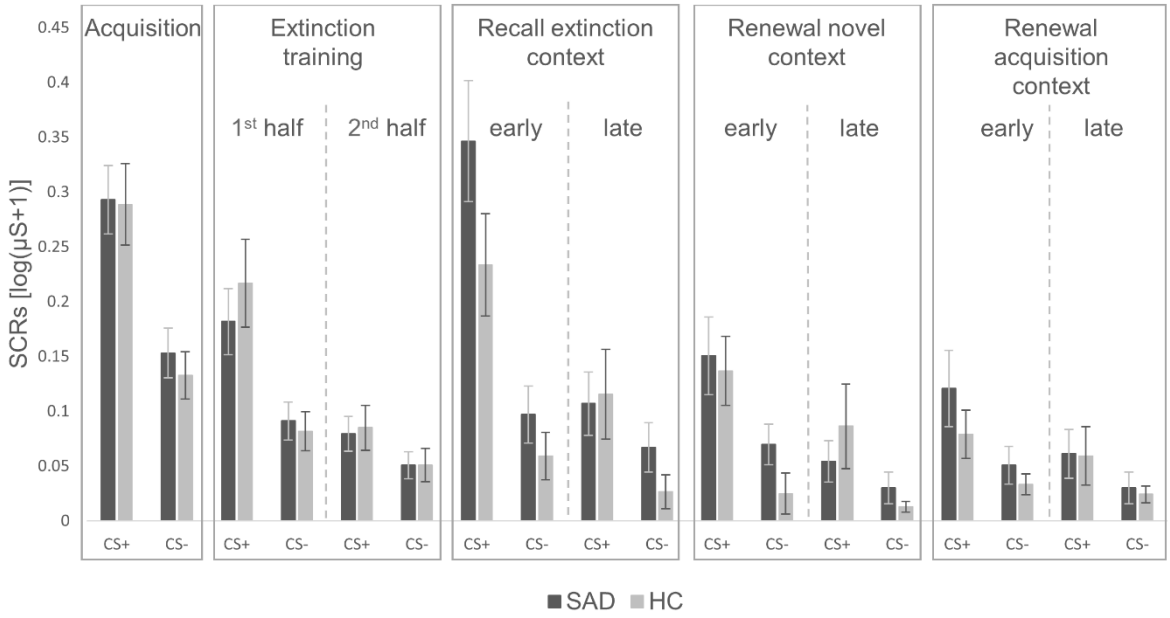

Figure 1. Skin conductance responses for CS+ and CS- during all phases of fear conditioning in patients with social anxiety disorder and healthy controls. SAD=social anxiety disorder; HC=healthy controls; µS=Microsiemens. '1st half' refers to the first eight trials, '2nd half' to the last eight trials of both CSs. The last two trials of each CS of the extinction training, which were used for the analyses on spontaneous recovery in the extinction context, are not illustrated here. 'Early' refers to the first four trials, 'late' to the last two trials of both CSs. Error bars represent standard errors of the mean.

Analyses of post hoc context ratings did not reveal a significant difference in the overall context recognition score ($t_{(107)}$=0.91, p=.364), indicating that both groups were equally accurate in telling which contexts they saw and when, and whether they were electrically stimulated in that context. A significant main effect of group was found for fear ratings ($F_{(1, 99)}$=9.19, p=.003), with the SAD group reporting more fear in all three contexts (acquisition context: $t_{(95.3)}$=.286, p=.005; extinction context: $t_{(100)}$=2.65, p=.009; novel context: $t_{(99)}$=2.54, p=.013; see Supplementary Figure 5). There were no significant main effects of or interaction effects with group regarding arousal, valence and UCS expectancy ratings (all p≥.073).

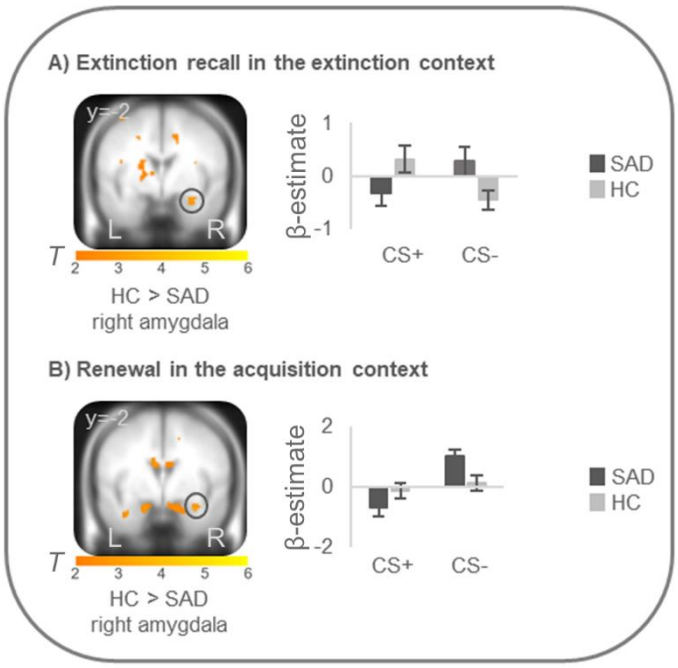

Figure 2. Stronger neural activation for differential conditioned responding (CS+ minus CS-) in the right amygdala in healthy controls in comparison to patients with social anxiety disorder during (A) extinction recall in the extinction context and (B) renewal in the acquisition context. Social anxiety disorder=SAD, healthy controls=HC. Figure shows stronger differential conditioned activation in the right amygdala in HC than in SAD for A recall in the extinction context (T_{max}=3.52, p_{FWE}=.024, MNI: x=30, y=-2, z=-22) and for B renewal in the acquisition context (T_{max}=3.31, p_{FWE}=.039, MNI: x=32, y=-2, z=-22). Coordinates are given in MNI space; color bars depict T-values; activations were superimposed on the MNI305 T1 template; bar graphs depict contrast values in peak voxels for CS+ and CS- separately; error bars depict standard errors of the mean.

Comparisons of patients with clinically relevant and without intrusion symptoms

INT did not differ from NO-INT in age, sex, and years of education, nor in severity of social anxiety symptoms (see Table 2). The INT group reported more intrusion symptoms regarding their aversive social experience and a higher depression score (BDI-II), but they did not meet a depression diagnosis more often than the NO-INT group.

There were no significant differences between the INT and the NO-INT group regarding SCRs during all conditioning phases (all $p \geq .053$; see Supplementary Figure 6).

On the neural level, differential conditioned responses did not differ between groups during fear acquisition. During extinction learning (first minus second half of extinction learning), right amygdala activation (CS+>CS-) more strongly decreased in the INT group compared with the NO-INT group (T_{max}=4.00, $p_{FWEcorr}$=.011, MNI: x=26, y=-4, z=-22). Connectivity analyses with this right amygdala activation peak as seed region showed a stronger connectivity with the right insula (T_{max}=4.73, $p_{FWEcorr}$=.005, MNI: x=38, y=14, z=-14) in the INT vs. NO-INT group. Stronger left vmPFC activation (CS+>CS-) was found for the NO-INT compared with the INT group, during the

late phase of extinction learning (T_{max}=3.89, $p_{FWEcorr}$=.049, MNI: x=-10, y=48, z=-18), as well as during early extinction recall in context B (T_{max}=4.23, $p_{FWEcorr}$=.021, MNI: x=-8, y=22, z=-24) and early renewal in the novel context C (T_{max}=4.42, $p_{FWEcorr}$=.014, MNI: x=-8, y=38, z=-10). The latter two activation differences are located in the posterior part of the vmPFC. In the novel context, PPI analyses revealed stronger connectivity of this left vmPFC activation locus with other subregions of bilateral vmPFC (T_{max}=4.01, $p_{FWEcorr}$=.041, MNI: x=-2, y=26, z=-14; T_{max}=4.29, $p_{FWEcorr}$=.02, MNI: x=0, y=26, z=-12) for the NO-INT vs. INT group. PPI analyses for vmPFC activation loci during late extinction learning and early extinction recall did not reveal any significant group differences. Since the sphere of the seed region (6mm radius) for the left vmPFC voxel regarding extinction recall in the extinction context was not fully located in the brain in all subjects, the sphere was reduced to a radius of 4mm for this single analysis, resulting in n=18 individuals in the INT and n=27 individuals in the NO-INT group. During renewal in the acquisition context A, the right amygdala (T_{max}=3.42, $p_{FWEcorr}$=.041, MNI: x=16, y=-8, z=-14) was activated more strongly in the INT compared with the NO-INT group (see Figure 3), but no significant group differences appeared for PPI analyses with this right amygdala activation locus. Because the INT and NO-INT groups differed in their BDI-II scores, we performed the analyses on the significant activation differences described above again and added BDI-II scores as a covariate of no interest. The group differences remain predominantly the same (see Supplement G).

Moreover, there were no significant differences between the INT and NO-INT group regarding the context recognition score ($t_{(41)}$=1.04, p=.303), nor UCS expectancy, fear, arousal, and valence ratings of the different contexts (all p≥.082).

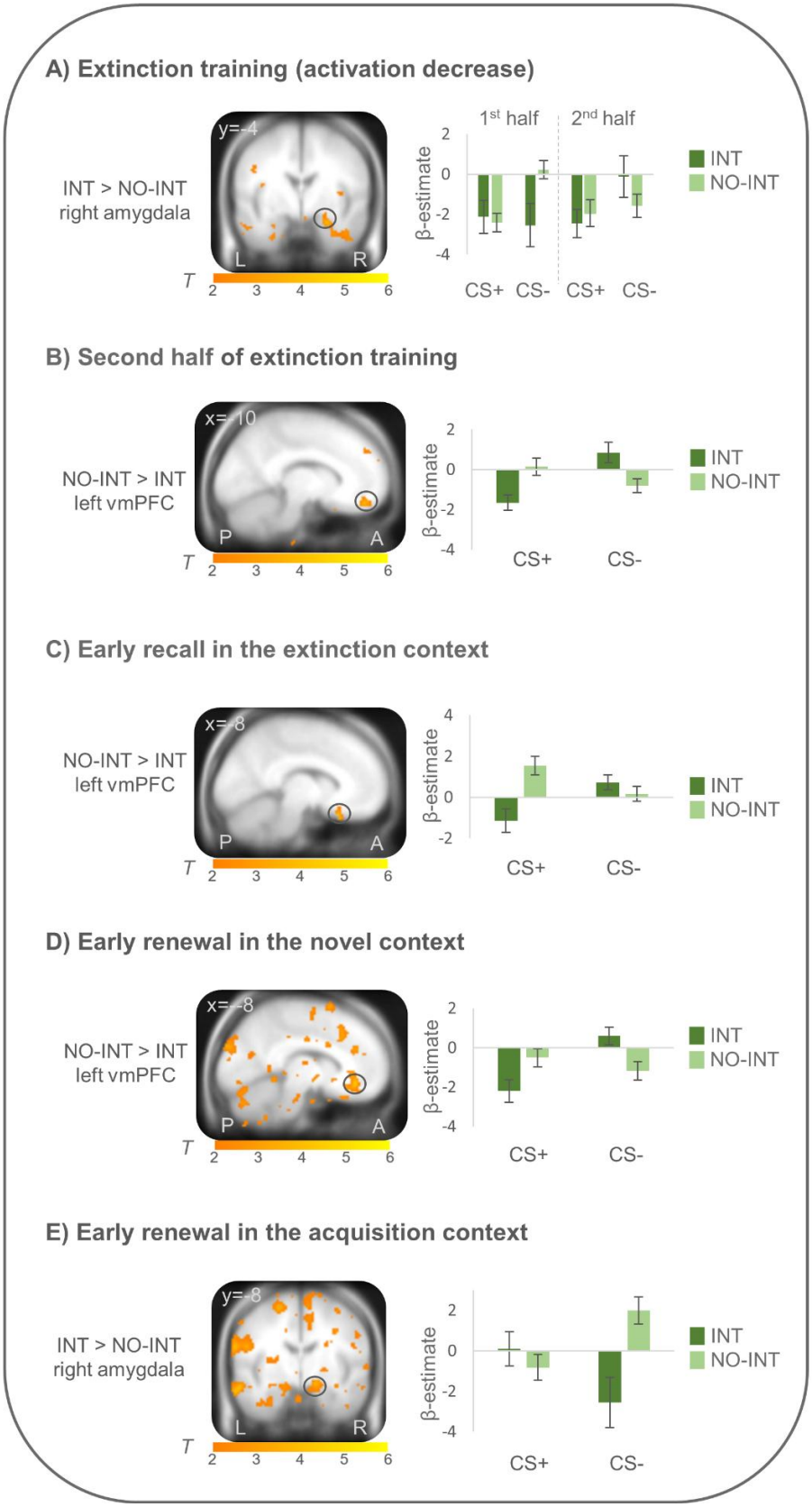

Figure 3. Neural activation of differential conditioned responding (CS+ minus CS-) during extinction training and extinction retrieval (~24h later) in patients with clinically relevant intrusions in comparison

to patients without intrusions. INT=patients with clinically relevant intrusions; NO-INT=patients without intrusions, '2nd half' refers to the last eight trials of both CSs; 'early' refers to the first four trials of both CSs. Figure shows **A)** significantly stronger decrease in right amygdala activation during extinction training (first minus second half, T_{max}=4.00, $p_{FWEcorr}$=.01, MNI: x=26, y=-4, z=-22), **B)** reduced vmPFC activation during late extinction (T_{max}=3.89, $p_{FWEcorr}$<.05, MNI: x=-10, y=48, z=-18), **C)** reduced vmPFC activation during early recall in the extinction context (T_{max}=4.23, $p_{FWEcorr}$=.02, MNI: x=-8, y=22, z=-24) and **D)** early renewal in the novel context (T_{max}=4.42, $p_{FWEcorr}$=.01, MNI: x=-8, y=38, z=-10), as well as **E)** stronger right amygdala activation during early renewal in the acquisition context (T_{max}=3.42, $p_{FWEcorr}$=.04, MNI: x=16, y=-8, z=-14) in patients with clinically relevant intrusions compared to patients without intrusions. Coordinates are given in MNI space; color bars depict T-values; activations were superimposed on the MNI305 T1 template; bar graphs depict contrast values in peak voxels for CS+ and CS- separately; error bars depict standard errors of the mean.

Discussion

The purpose of this study was to investigate neural correlates of context-dependent fear conditioning processes in SAD and their association with intrusion symptoms in response to an autobiographical aversive social event. In general, SAD patients showed reduced amygdala activation (CS+ vs. CS-) during extinction recall in the safe extinction context and during fear renewal in the acquisition context as well as stronger fear ratings for all three contexts. Moreover, a subgroup of patients with clinically relevant intrusions in response to the aversive social event (INT group) showed reduced activation of the vmPFC (CS+ vs. CS-) during extinction learning, extinction recall in the safe extinction context, and fear renewal in a novel context compared to a subgroup of patients without intrusions (NO-INT group), while there were no significant differences for self-report and SCRs.

Patients with SAD reported significantly more intrusions in response to aversive social events compared with HC. This finding is in line with previous results (Bjornsson et al., 2020; Erwin et al., 2006) and underpins the relevance of aversive social experiences and related PTSS in SAD. Regarding the context-dependent fear conditioning paradigm, reduced activation of the right amygdala in SAD compared to HC extends previous results showing diminished vmPFC activation during (context-dependent) extinction recall in social anxiety (disorder) (Marin et al., 2020; Pejic et al., 2013). In part, our findings correspond to results of a previous study in PTSD, showing reduced SCRs and amygdala activation towards the extinguished CS+ during fear renewal in the

acquisition context, which might indicate an inability to use contextual information to regulate conditioned fear (Garfinkel et al., 2014). However, visual inspection of the results suggests that the activation differences observed in our study are not due to a stronger recall of fear in HC, but more likely result from a stronger activation towards the CS- in SAD. In line with this, patients with SAD compared with HC reported more fear in all three contexts, including the safe extinction context. Thus, these findings might indicate poorer safety learning in SAD, which is consistent with previous literature (Chauret et al., 2019; Duits et al., 2015; Savage, Davey, Fullana, & Harrison, 2020) and has also been shown in PTSD (Garfinkel et al., 2014; Jovanovic, Kazama, Bachevalier, & Davis, 2012). However, other studies found a positive association between symptoms of SAD and safety reversal learning (Savage et al., 2020). Nevertheless, the meaning of activation/deactivation of the amygdala has not been fully elucidated yet. Previous studies showed inconsistent results regarding the role of the amygdala in extinction learning (Fullana et al., 2018), with different subnuclei of the amygdala and cell types within the amygdala probably subserving different functions in associative fear and extinction processes (Likhtik, Popa, Apergis-Schoute, Fidacaro, & Paré, 2008; Repa et al., 2001). Moreover, amygdala deactivation may also be related to other processes as e.g. emotion regulation (Frank et al., 2014) or coping (Petrovic, Carlsson, Petersson, Hansson, & Ingvar, 2004).

Differences between our and previous findings regarding fear conditioning processes in SAD might be due to methodological issues as e.g. the use of different stimuli (e.g. disorder specific vs. unspecific stimuli) (Chauret et al., 2019; Lissek et al., 2008; Pejic et al., 2013), different outcome measures or differences in sample size. All in all, our results do not show a clear correspondence between abnormalities in patients with SAD and PTSD regarding context-dependent fear conditioning.

Regarding the association of intrusion symptoms and context-dependent fear conditioning, the INT compared with the NO-INT group showed reduced vmPFC activation during late extinction training, early extinction recall, and early renewal in the novel context. During renewal in the novel

context, the INT group also exhibited less connectivity within the vmPFC. These findings fit previous literature on vmPFC hypoactivation during extinction learning as well as extinction recall in PTSD (Diekhof, Geier, Falkai, & Gruber, 2011; Garfinkel et al., 2014; Milad et al., 2009; Rougemont-Bücking et al., 2011; Shvil, Rusch, Sullivan, & Neria, 2013; Suarez-Jimenez et al., 2020; VanElzakker et al., 2018) and also extinction recall in SAD (Marin et al., 2020). Since activation of the vmPFC is associated with safety or extinction learning (Fullana et al., 2016; Milad et al., 2007; Phelps, Delgado, Nearing, & LeDoux, 2004), these results support our hypothesis of a link between dysfunctional regulation of fear and intrusion symptoms. Another study (Garfinkel et al., 2014) also found lower activation of the vmPFC in PTSD during renewal (in response to the extinguished CS+), but in the acquisition context, and at the same time lower SCRs. These results suggest that vmPFC activation might not have a fear-inhibiting effect per se, but rather a fear-modulating effect. Missing differences in conditioned SCRs and subjective responses as well as vmPFC activation differences in the novel but not the acquisition context in our study could, however, not fully support this interpretation. While it is well known that the vmPFC is involved in the inhibition of fear expression via connections to the amygdala (Motzkin, Philippi, Wolf, Baskaya, & Koenigs, 2015; Quirk & Mueller, 2008), the abovementioned findings suggest that the role of the vmPFC in fear conditioning processes is more complex. Moreover, it is involved in the acquisition of new learning content (Battaglia, Garofalo, Di Pellegrino, & Starita, 2020), or the integration of information (Nieuwenhuis & Takashima, 2011), e.g. by using contextual information to regulate fear (Gonzalez & Fanselow, 2020; Hermann et al., 2016; Kalisch et al., 2006; Pennington, Anderson, & Fanselow, 2017). Furthermore, previous studies have identified subregions of the vmPFC associated with different states (anticipated threat or processing of safety signals) (Battaglia, Harrison, & Fullana, 2022; Harrison et al., 2017; Tashjian, Zbozinek, & Mobbs, 2021). Considering different activation loci within the vmPFC as found in our study, a more detailed analyses of vmPFC subregions could help to more clearly understand its role in (context-dependent) fear conditioning.

Moreover, the INT compared to the NO-INT group exhibited altered activation of the right amygdala: during extinction training they showed a greater decrease of conditioned amygdala activation which could be interpreted as an indicator for both, a stronger reduction of conditioned fear or deficient extinction learning regarding the abovementioned different functions of amygdala subregions during extinction learning (Likhtik et al., 2008; Repa et al., 2001). In addition, this result is contrary to previous findings in PTSD showing reduced decrease of amygdala activation during extinction learning (Milad et al., 2009; Suarez-Jimenez et al., 2020). In a study with healthy participants, social anxiety was negatively correlated with amygdala activity during extinction learning, which may indicate reduced extinction learning in people with stronger social anxiety (Pejic et al., 2013). Moreover, hyperactivation of the right amygdala during renewal in the acquisition context in the INT compared with the NO-INT group was not in line with a previous study, where patients with PTSD demonstrated attenuated fear renewal in SCRs and decreased amygdala activity, however only in response to the CS+ alone (Garfinkel et al., 2014). Visual inspection of the data also indicates that primarily activation differences in response to the CS- contributed to the abovementioned amygdala findings. All in all, our findings regarding altered amygdala activation in INT vs. NO-INT patients are difficult to interpret and need further investigation.

A limitation of our study is that main effects analyses in both groups (SAD and HC) revealed a significant return of fear during extinction recall in the extinction context on the second day, which might be due to immediate extinction training after fear acquisition without a consolidation phase (Merz, Hamacher-Dang, & Wolf, 2016). Moreover, the specific instruction that contingencies will not change during the fear conditioning paradigm might have contributed to this effect. In addition, the INT group exhibited more severe depressive symptoms compared with the NO-INT group. Besides intrusion symptoms, this factor might also have contributed to deficient context-dependent modulation of conditioned fear.

Our study also raises a methodological discussion point. Multiple outcome measures ideally allow for a broad and clear interpretation of results. However, it also raises the question of the relevance of various biological measures, such as neural activation and SCR, if these are not reflected in subjective outcome measures (e.g., fear rating), or vice versa (see e.g. Ahrens et al., 2016).

Our results provide evidence for the transdiagnostic relevance of intrusion symptoms in association with context-dependent fear conditioning processes in SAD. Regarding implications for the treatment of SAD, it should be addressed how exposure therapy in SAD can be designed to facilitate recall of extinction memories in different contexts. Imagery rescripting seems to be a promising intervention in order to overcome aversive social memories in SAD (Norton & Abbott, 2017; Reimer & Moscovitch, 2015) and its effects on extinction recall and fear renewal in real-life social situations should be investigated in futures studies.

References

Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of Anxiety Disorders*, *44*, 36–46. https://doi.org/10.1016/j.janxdis.2016.10.003

Battaglia, S., Garofalo, S., Di Pellegrino, G., & Starita, F. (2020). Revaluing the Role of vmPFC in the Acquisition of Pavlovian Threat Conditioning in Humans. *The Journal of Neuroscience*, *40*(44), 8491–8500. https://doi.org/10.1523/JNEUROSCI.0304-20.2020

Battaglia, S., Harrison, B. J., & Fullana, M. A. (2022). Does the human ventromedial prefrontal cortex support fear learning, fear extinction or both? A commentary on subregional contributions. *Molecular Psychiatry*, *27*(2), 784–786. https://doi.org/10.1038/s41380-021-01326-4

Bjornsson, A. S., Hardarson, J. P., Valdimarsdottir, A. G., Gudmundsdottir, K., Tryggvadottir, A., Thorarinsdottir, K., . . . Thorisdottir, A. S. (2020). Social trauma and its association with posttraumatic stress disorder and social anxiety disorder. *Journal of Anxiety Disorders*, *72*, 102228. https://doi.org/10.1016/j.janxdis.2020.102228

Bouton, M. E., Westbrook, R. F., Corcoran, K. A., & Maren, S. (2006). Contextual and temporal modulation of extinction: Behavioral and biological mechanisms. *Biological Psychiatry*, *60*(4), 352–360. https://doi.org/10.1016/j.biopsych.2005.12.015

Chauret, M., Suffren, S., Pine, D. S., Nassim, M., Saint-Amour, D., & Maheu, F. S. (2019). Fear conditioning and extinction in anxious youth, offspring at-risk for anxiety and healthy comparisons:

An fMRI study. *Biological Psychology*, *148*, 107744. https://doi.org/10.1016/j.biopsycho.2019.107744

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Foa, E. B., & Weisler, R. H. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *The British Journal of Psychiatry: The Journal of Mental Science*, *176*, 379–386. https://doi.org/10.1192/bjp.176.4.379

Consbruch, K., Stangier, U., & Heidenreich, T. (2016). *Soziale Angst-Skalen (SOZAS). SIAS, SPS, SPIN, LSAS*. Göttingen: Hogrefe.

Cwik, J. C., & Woud, M. L. (2015). Deutsche Übersetzung der CLINICIAN-ADMINISTERED PTSD SCALE for DSM-5 (CAPS-5). *Health Research and Treatment Center*.

Diekhof, E. K., Geier, K., Falkai, P., & Gruber, O. (2011). Fear is only as deep as the mind allows: A coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *NeuroImage*, *58*(1), 275–285. https://doi.org/10.1016/j.neuroimage.2011.05.073

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., . . . Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, *32*(4), 239–253. https://doi.org/10.1002/da.22353

Dvir, M., Horovitz, O., Aderka, I. M., & Shechner, T. (2019). Fear conditioning and extinction in anxious and non-anxious youth: A meta-analysis. *Behaviour Research and Therapy*, *120*, 103431. https://doi.org/10.1016/j.brat.2019.103431

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

Erwin, B. A., Heimberg, R. G., Marx, B. P., & Franklin, M. E. (2006). Traumatic and socially stressful life events among persons with social anxiety disorder. *Journal of Anxiety Disorders*, *20*(7), 896–914. https://doi.org/10.1016/j.janxdis.2005.05.006

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*(2), 175–191. https://doi.org/10.3758/BF03193146

Frank, D. W., Dewitt, M., Hudgens-Haney, M., Schaeffer, D. J., Ball, B. H., Schwarz, N. F., . . . Sabatinelli, D. (2014). Emotion regulation: Quantitative meta-analysis of functional activation and deactivation. *Neuroscience and Biobehavioral Reviews*, *45*, 202–211. https://doi.org/10.1016/j.neubiorev.2014.06.010

Franke, L. K., Rattel, J. A., Miedl, S. F., Danböck, S. K., Bürkner, P.-C., & Wilhelm, F. H. (2021). Intrusive memories as conditioned responses to trauma cues: An empirically supported concept? *Behaviour Research and Therapy*, *143*, 103848. https://doi.org/10.1016/j.brat.2021.103848

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine*, *31*(6), 1025–1035. https://doi.org/10.1017/S0033291701004056

Fullana, M. A., Albajes-Eizagirre, A., Soriano-Mas, C., Vervliet, B., Cardoner, N., Benet, O., . . . Harrison, B. J. (2018). Fear extinction in the human brain: A meta-analysis of fMRI studies in healthy participants. *Neuroscience and Biobehavioral Reviews*, *88*, 16–25. https://doi.org/10.1016/j.neubiorev.2018.03.002

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

Garfinkel, S. N., Abelson, J. L., King, A. P., Sripada, R. K., Wang, X., Gaines, L. M., & Liberzon, I. (2014). Impaired contextual modulation of memories in PTSD: An fMRI and psychophysiological study of extinction retention and fear renewal. *The Journal of Neuroscience*, *34*(40), 13435–13443. https://doi.org/10.1523/JNEUROSCI.4287-13.2014

Gonzalez, S. T., & Fanselow, M. S. (2020). The role of the ventromedial prefrontal cortex and context in regulating fear learning and extinction. *Psychology & Neuroscience*, *13*(3), 459–472. https://doi.org/10.1037/pne0000207

Hackmann, A., Clark, D. M., & McManus, F. (2000). Recurrent images and early memories in social phobia. *Behaviour Research and Therapy*, *38*(6), 601–610. https://doi.org/10.1016/S0005-7967(99)00161-8

Harrison, B. J., Fullana, M. A., Via, E., Soriano-Mas, C., Vervliet, B., Martínez-Zalacaín, I., . . . Cardoner, N. (2017). Human ventromedial prefrontal cortex and the positive affective processing of safety signals. *NeuroImage*, *152*, 12–18. https://doi.org/10.1016/j.neuroimage.2017.02.080

Hermann, A., Stark, R., Milad, M. R., & Merz, C. J. (2016). Renewal of conditioned fear in a novel context is associated with hippocampal activation and connectivity. *Social Cognitive and Affective Neuroscience*, *11*(9), 1411–1421. https://doi.org/10.1093/scan/nsw047

Hermann, A., Stark, R., Müller, E. A., Kruse, O., Wolf, O. T., & Merz, C. J. (2020). Multiple extinction contexts modulate the neural correlates of context-dependent extinction learning and retrieval. *Neurobiology of Learning and Memory*, *168*, 107150. https://doi.org/10.1016/j.nlm.2019.107150

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

Kalisch, R., Korenfeld, E., Stephan, K. E., Weiskopf, N., Seymour, B., & Dolan, R. J. (2006). Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. *The Journal of Neuroscience*, *26*(37), 9503–9511. https://doi.org/10.1523/JNEUROSCI.2021-06.2006

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, *22*, 141–173. https://doi.org/10.1159/000414022

Likhtik, E., Popa, D., Apergis-Schoute, J., Fidacaro, G. A., & Paré, D. (2008). Amygdala intercalated neurons are required for expression of fear extinction. *Nature*, *454*(7204), 642–645. https://doi.org/10.1038/nature07167

Lissek, S., & Grillon, C. (2012). Learning Models of PTSD. In J. G. Beck & D. M. Sloan (Eds.), *Oxford library of psychology. The Oxford Handbook of Traumatic Stress Disorders* (pp. 174–190). Oxford, New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399066.013.0013

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *The American Journal of Psychiatry*, *165*(1), 124–132. https://doi.org/10.1176/appi.ajp.2007.06091513

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

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews. Neuroscience*, *14*(6), 417–428. https://doi.org/10.1038/nrn3492

Margraf, J., Cwik, J. C., Suppiger, A., & Schneider, S. (2017). *DIPS Open Access 1.2: Diagnostisches Interview bei psychischen Störungen*. Ruhr-Universität Bochum (RUB). https://doi.org/10.13154/RUB.100.89

Marin, M.-F., Hammoud, M. Z., Klumpp, H., Simon, N. M., & Milad, M. R. (2020). Multimodal Categorical and Dimensional Approaches to Understanding Threat Conditioning and Its Extinction in Individuals With Anxiety Disorders. *JAMA Psychiatry*, *77*(6), 618–627. https://doi.org/10.1001/jamapsychiatry.2019.4833

Marin, M.-F., Zsido, R. G., Song, H., Lasko, N. B., Killgore, W. D. S., Rauch, S. L., . . . Milad, M. R. (2017). Skin Conductance Responses and Neural Activations During Fear Conditioning and Extinction Recall Across Anxiety Disorders. *JAMA Psychiatry*, *74*(6), 622–631. https://doi.org/10.1001/jamapsychiatry.2017.0329

McCabe, R. E., Antony, M. M., Summerfeldt, L. J., Liss, A., & Swinson, R. P. (2003). Preliminary examination of the relationship between anxiety disorders in adults and self-reported history of teasing or bullying experiences. *Cognitive Behaviour Therapy*, *32*(4), 187–193. https://doi.org/10.1080/16506070310005051

Merz, C. J., Hamacher-Dang, T. C., & Wolf, O. T. (2016). Immediate extinction promotes the return of fear. *Neurobiology of Learning and Memory*, *131*, 109–116. https://doi.org/10.1016/j.nlm.2016.03.013

Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., . . . Rauch, S. L. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. *Biological Psychiatry*, *66*(12), 1075–1082. https://doi.org/10.1016/j.biopsych.2009.06.026

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry*, *62*(5), 446–454. https://doi.org/10.1016/j.biopsych.2006.10.011

Mineka, S., & Zinbarg, R. E. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It's not what you thought it was. *The American Psychologist*, *61*(1), 10–26. https://doi.org/10.1037/0003-066X.61.1.10

Motzkin, J. C., Philippi, C. L., Wolf, R. C., Baskaya, M. K., & Koenigs, M. (2015). Ventromedial prefrontal cortex is critical for the regulation of amygdala activity in humans. *Biological Psychiatry*, *77*(3), 276–284. https://doi.org/10.1016/j.biopsych.2014.02.014

Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research*, *218*(2), 325–334. https://doi.org/10.1016/j.bbr.2010.12.009

Norton, A. R., & Abbott, M. J. (2017). Bridging the Gap between Aetiological and Maintaining Factors in Social Anxiety Disorder: The Impact of Socially Traumatic Experiences on Beliefs, Imagery and Symptomatology. *Clinical Psychology & Psychotherapy*, *24*(3), 747–765. https://doi.org/10.1002/cpp.2044

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

Pennington, Z. T., Anderson, A. S., & Fanselow, M. S. (2017). The ventromedial prefrontal cortex in a model of traumatic stress: Fear inhibition or contextual processing? *Learning & Memory*, *24*(9), 400–406. https://doi.org/10.1101/lm.046110.117

Petrovic, P., Carlsson, K., Petersson, K. M., Hansson, P., & Ingvar, M. (2004). Context-dependent deactivation of the amygdala during pain. *Journal of Cognitive Neuroscience*, *16*(7), 1289–1301. https://doi.org/10.1162/0898929041920469

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron*, *43*(6), 897–905. https://doi.org/10.1016/j.neuron.2004.08.042

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

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology*, *46*(5), 984–995. https://doi.org/10.1111/j.1469-8986.2009.00852.x

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, *33*(1), 56–72. https://doi.org/10.1038/sj.npp.1301555

Reimer, S. G., & Moscovitch, D. A. (2015). The impact of imagery rescripting on memory appraisals and core beliefs in social anxiety disorder. *Behaviour Research and Therapy*, *75*, 48–59. https://doi.org/10.1016/j.brat.2015.10.007

Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., & LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nature Neuroscience*, *4*(7), 724–731. https://doi.org/10.1038/89512

Rougemont-Bücking, A., Linnman, C., Zeffiro, T. A., Zeidan, M. A., Lebron-Milad, K., Rodriguez-Romaguera, J., . . . Milad, M. R. (2011). Altered processing of contextual information during fear

extinction in PTSD: An fMRI study. *CNS Neuroscience & Therapeutics*, *17*(4), 227–236. https://doi.org/10.1111/j.1755-5949.2010.00152.x

Savage, H. S., Davey, C. G., Fullana, M. A., & Harrison, B. J. (2020). Threat and safety reversal learning in social anxiety disorder - an fMRI study. *Journal of Anxiety Disorders*, *76*, 102321. https://doi.org/10.1016/j.janxdis.2020.102321

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., & Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: A systematic review. *PLOS ONE*, *4*(6), e5865. https://doi.org/10.1371/journal.pone.0005865

Shvil, E., Rusch, H. L., Sullivan, G. M., & Neria, Y. (2013). Neural, psychophysiological, and behavioral markers of fear processing in PTSD: A review of the literature. *Current Psychiatry Reports*, *15*(5), 358. https://doi.org/10.1007/s11920-013-0358-3

Sosic, Z., Gieler, U., & Stangier, U. (2008). Screening for social phobia in medical in- and outpatients with the German version of the Social Phobia Inventory (SPIN). *Journal of Anxiety Disorders*, *22*(5), 849–859. https://doi.org/10.1016/j.janxdis.2007.08.011

Spies, J.-P., Woud, M. L., Kessler, H., Rau, H., Willmund, G. D., Köhler, K., . . . Cwik, J. C. (2020). Psychometric properties of the German version of the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) in clinical routine settings: Study design and protocol of a multitrait-multimethod study. *BMJ Open*, *10*(6), e036078. https://doi.org/10.1136/bmjopen-2019-036078

Suarez-Jimenez, B., Albajes-Eizagirre, A., Lazarov, A., Zhu, X., Harrison, B. J., Radua, J., . . . Fullana, M. A. (2020). Neural signatures of conditioning, extinction learning, and extinction recall in posttraumatic stress disorder: A meta-analysis of functional magnetic resonance imaging studies. *Psychological Medicine*, *50*(9), 1442–1451. https://doi.org/10.1017/S0033291719001387

Tashjian, S. M., Zbozinek, T. D., & Mobbs, D. (2021). A Decision Architecture for Safety Computations. *Trends in Cognitive Sciences*, *25*(5), 342–354. https://doi.org/10.1016/j.tics.2021.01.013

VanElzakker, M. B., Staples-Bradley, L. K., & Shin, L. M. (2018). The Neurocircuitry of Fear and PTSD. In E. Vermetten, A. Germain, & T. C. Neylan (Eds.), *Sleep and Combat-Related Post Traumatic Stress Disorder* (Vol. 113, pp. 111–125). New York: Springer New York. https://doi.org/10.1007/978-1-4939-7148-0_10

Wicking, M., Steiger, F., Nees, F., Diener, S. J., Grimm, O., Ruttorf, M., . . . Flor, H. (2016). Deficient fear extinction memory in posttraumatic stress disorder. *Neurobiology of Learning and Memory*, *136*, 116–126. https://doi.org/10.1016/j.nlm.2016.09.016

Wong, Q. J. J., & Rapee, R. M. (2016). The aetiology and maintenance of social anxiety disorder: A synthesis of complimentary theoretical models and formulation of a new integrated model. *Journal of Affective Disorders*, *203*, 84–100. https://doi.org/10.1016/j.jad.2016.05.069

Supplemental Information

A) Course of the entire project

The whole project consisted of six appointments. Diagnostic interviews and questionnaires, as well as interviews on various autobiographical events took place on a first appointment. Neuropsychological testing and an experiment on pattern separation were conducted on a second appointment. The third and fourth sessions, at which the fMRI paradigm reported here was performed, took place within 14 days of the diagnostic interviews and followed each other for approximately 24 hours. At the second scanning session, after the post hoc ratings of the fear conditioning paradigm, we conducted an experiment on the processing of neutral faces. The patient sample was invited for two more appointments to examine both, memory features and posttraumatic stress symptoms related to the autobiographical events, as well as the effect of Imagery Rescripting (see Seinsche et al., 2023).

B) Context-dependent fear conditioning paradigm

Each trial lasted for a total of 20s and started (jittered between 0.625-2.5 s), as well as ended (jittered between 8.500-10.6875 s) with a white fixation cross on a black background. In the first session, fear acquisition took place in context A (e.g., desk). The lamp was first turned off for 3 seconds and then lit up blue or yellow for 6 seconds. There were 16 trials during which the lamp lit blue (8x) or yellow (8x). One of the two colors was followed by electrical stimulation (UCS) in 62.5 % of the trials (5 out of 8 trials), the other color was never followed by the UCS. Thus, the lamp color was used as a CS+ or CS-. The first and last two trials each contained one CS+ and one CS-. The CS categories were presented in a pseudorandomized order over the remaining trials, so that no more than two trials with the same CS followed each other; the orders were counterbalanced over the two groups (SAD and HC). The first CS+ was always followed by the UCS to enhance fear acquisition. Likewise, the last CS+ was reinforced to ensure that no extinction takes place at this phase. The subjects were informed that they might receive electrical stimulation after the lamp was turned on and that there might be a relationship between the color of the lamp and the electrical stimulation. After acquisition,

participants were asked (among others) whether they noticed any relationship between the light colors and electrical stimulation, and if so, what kind of association (correct association: SAD n=46, HC n=45, $\chi^2(1)$=1.92, p=.17; INT n=18, NO-INT n=22, $\chi^2(1)$=.02, p=.89). To ensure awareness of the correct CS-UCS relationship in all participants, they were informed about it after the fear acquisition phase.

Immediately after this interview, which took place via a microphone/speaker in the scanner, the extinction training in context B followed. The CS+ was never followed by the UCS in this phase. Participants were informed in advance that they could receive electrical stimulation, and the learned relationship of the light colors and following UCS would remain the same in this case. There were 32 trials in which the CS+/CS- was presented 16 times each and which were divided into two blocks. Within these two blocks, the frequency of CS+/CS- was kept the same. The order was pseudorandomized, where the same CS was never presented more than two times in a row and the first two trials contained one CS+ and one CS-.

Approximately 24 hours later, firstly extinction recall in context B, secondly renewal in a novel context C, and thirdly fear renewal in the acquisition context A took place, each with 16 trials (divided into two blocks of eight trials with four CS+/four CS- each). CS+/CS- order was pseudorandomized, with again the same CS presented no more than twice in a row and the first two trials containing one CS+ and one CS-. Participants were informed, that they could receive electrical stimulation. If this was the case, the relationship (CS-UCS) would remain the same as learned during acquisition and the intensity would be the same as the previous day. The participants did not receive any electrical stimulation on the second day.

C) Post hoc rating

With respect to six different contexts (acquisition, extinction, novel and 3 unknown contexts), subjects were asked to indicate whether they had seen it 'only yesterday', 'today and yesterday' (acquisition and extinction context), 'only today' (novel context), or 'never'. They were also asked to indicate whether the UCS was presented in the particular contexts. With this information, recognition scores were calculated for the six different contexts (correct

indication of when seen and whether UCS was presented=1, otherwise=0), as well as an overall sum score for correct context assignment. Subsequently, they were asked (on a 9-point scale from 'sure that electrical stimulation will not follow' to 'sure that electrical stimulation will follow') how sure they were about the probability of receiving electrical stimulation in this context during the experiment. Participants were also asked to indicate valence, arousal, and fear in relation to the turned-off (no CS) and turned-on (blue/yellow) lamp in each context on 9-point scales ('very unpleasant' to 'very pleasant', 'calm and relaxed' to 'very aroused', 'not anxious at all' to 'very anxious').

D) Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging was conducted with a 3 Tesla whole body scanner with a 64 channel standard head coil (Siemens Prisma, Erlangen, Germany). For the normalization procedure, T1-weighted sagittal images (MPRAGE: 176 slices, flip-angle=8°, TI=900 ms, TR=1.58 s, 0.9375x0.9375x0.94 mm voxel size, acceleration: iPAT 3 [GRAPPA], duration=4:29 minute) were acquired. Three T2*-weighted gradient echo-planar imaging sequence (EPI) with 42 slices (slice thickness=3 mm; 0.6 mm gap; descending slice order; TE=30 ms; TR=2.5 s; flip angle=81°; Field of View=220x220 mm²; matrix size=74x74 pixel, parallel imaging: GRAPPA acceleration 2; oblique slices 30 degrees tilt from the AC-PC line) covering the entire brain was used to assess BOLD responses. The functional image acquisition consisted of 152 volumes (6:20 minutes) for fear acquisition, 280 volumes (11:40 minutes) for extinction training as well as 408 volumes (17 minutes) for extinction recall and renewal in contexts C and A.

To correct for geometric distortions of echo-planar images, whole-brain gradient echo based fieldmaps (slice thickness: 3 mm, 0.6 mm gap, flip-angle: 90°, echo times: 10.0 and 12.46 ms, FoV: 220x220 mm) were acquired in each session.

We used Statistical Parametric Mapping (SPM12, r7219, Wellcome Department of Cognitive Neurology, London, UK; 2009), implemented in Matlab R2019a (Mathworks Inc., Sherborn, MA, USA) for data preprocessing and analysis.

Functional brain imaging data were corrected for motion and geometric distortion by means of the joint realign and unwarp procedure implemented in SPM (registered to first image, interpolation 4th degree b-spline). Afterwards, images were slice-time corrected to the middle slice (Fourier shift). Individual high-resolution structural scans were segmented using the unified segmentation approach (incl. bias correction) to obtain deformation fields for normalization to MNI space (12 parameter affine + deformation field, ICBM space template, European brains, tissue probability maps, unified segmentation). Slice time corrected functional images were coregistered to the structural scan (rigid body, mutual information) and were normalized using the deformation fields (resulting voxel size: 2x2x2 mm, 4th degree b-spline). In a last step functional images were smoothed by convolution of an isotropic three-dimensional Gaussian kernel of 6 mm full-width at half maximum (FWHM).

Separate models for first level analyses were set up for fear acquisition, extinction training and retrieval testing (extinction recall and fear renewal in the novel and the acquisition context). Regressors in each model were: context alone, blocks of eight (acquisition and extinction training) or four (retrieval testing) trials for CS+ and CS- for an early and a late phase, respectively. In the acquisition model, there were additional regressors for UCS as well as noUCS time points. As regressors of no interest we added the six movement parameters from realignment and one regressor for each volume with a framewise displacement >0.5 mm (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012). All regressors of interest were modeled based on a stick function convolved with the canonical hemodynamic response function (HRF) in the general linear model, without any specific modeling of the duration of the distinct events (i.e., event-based design). Autocorrelation of time series was controlled by AR(1) + white noise model and a 128 s high-pass filter (via discrete cosine transform).

Region of Interest (ROI) analyses were done for amygdala, insula, hippocampus, (probability masks taken from the current ‘Harvard-Oxford Cortical and Subcortical Structural Atlases`, Harvard Center for Morphometric Analysis [http://www.cma.mgh.harvard.edu/], probability threshold: 0.50) included in the FSL software package (http://www.fmrib.ox.ac.uk/fsl/). Because there are no masks in these atlases that specifically

map the dACC and vmPFC, we constructed them with the MARINA software package (Walter, 2002). The vmPFC mask is composed of the bilateral medial orbital area of the frontal cortex and the gyrus rectus according to the parcellation of Tzourio-Mazoyer et al. (2002). The mask for the dACC was constructed as the anterior and middle cingulate cortex from the AAL parcellation (Tzourio-Mazoyer et al., 2002) ranging from y=-17 mm to y=32 mm, z≥12 mm (MNI coordinates). These masks have been used in previous studies (e.g. Hermann et al., 2020; Sperl et al., 2019; see Supplementary Figure 7 for an illustration of the masks used). We used the small volume correction option of SPM12 and a significance threshold of α≤.05 on voxel level with family-wise error (FWE) correction. The reported coordinates (x, y, z) refer to the MNI space.

E) Skin Conductance Responses

Data were recorded with a sampling rate of 1000 Hz. The raw data were down sampled to 100 Hz and smoothed with a Gaussian Kernel of 32 mm. Data quality was checked manually and technical artifacts were corrected via manual interpolation, which is integrated in Ledalab. For the 'trough-to-peak' analyses (Pineles, Orr, & Orr, 2009) we defined the entire interval response (EIR) for the CS within a time window of 0.8-6 s after CS onset and for the UCS within a time window of 0.8-2.5 s after UCS onset.

F) Main effects of fear conditioning across groups

Across both groups, stronger SCRs (main effect stimulus, $F_{(1, 90)}$=77.09, p<.001) as well as stronger BOLD signals in bilateral amygdala, dACC, hippocampus, insula, and vmPFC were observed for the CS+ compared to the CS-, reflecting successful fear acquisition. In contrast, activity in the left vmPFC was reduced (stronger BOLD signals for the CS- compared to the CS+; see Supplementary Table 4 for all neural activation differences in the phases of the conditioning paradigm).

Furthermore, SCR data indicate successful extinction learning, as reflected by a significant decrease from early to late extinction training for CS+ vs. CS- (interaction effect between stimulus and phase, Greenhouse-Geisser $F_{(1, 90)}$=20.24, p<.001 with significant less differential conditioned SCRs in the second half than in the first half of extinction training) as

well as no more differences between CS+ and CS- during the last two trials (main effect stimulus, $F_{(1, 90)}$=2.22, p=.14). On the neural level, there was a significant decrease in activity (CS+>CS-) in the left vmPFC from early to late extinction training. During early extinction, stronger differential conditioned responding (CS+>CS-) was found in bilateral insula and dACC, as well as diminished activation (CS->CS+) in right amygdala, left hippocampus, and right vmPFC. In the second half of extinction, a stronger differential conditioned response (CS+>CS-) appeared in the left insula, while right insula, bilateral amygdala and bilateral vmPFC activation was reduced (CS->CS+).

Stronger differential conditioned SCRs during retrieval testing in the extinction context one day later compared to late extinction training (interaction effect between stimulus and phase, Greenhouse Geiser $F_{(1, 90)}$=27.59, p<.001) indicate a spontaneous recovery of the conditioned response. There was a significant increase in SCRs towards the CS+ from late extinction training to early retrieval testing ($t_{(91)}$=-5.75, p>.001) while there was no significant increase for the CS- ($t_{(91)}$=-1.71, p=.09). This is supported by a significant differential activation (CS+>CS-) of brain areas belonging to the fear network (bilateral dACC and insula) during extinction recall in the safe extinction context. Moreover, reduced activation (CS->CS+) in bilateral amygdala, hippocampus and vmPFC was found.

Early fear renewal in the novel and the acquisition context was reflected by stronger differential conditioned SCRs (main effect stimulus, novel context: $F_{(1, 90)}$=25.59, p<.001 and acquisition context: $F_{(1, 90)}$=9.99, p=.002) as well as a significant increase of SCRs from late renewal in the novel context to early renewal in the acquisition context (main effect phase, Greenhouse-Geiser $F_{(1, 90)}$=4.86, p=.03), but not from late extinction recall to early renewal in the novel context (main effect phase, Greenhouse-Geiser $F_{(1, 90)}$=1.69, p=.197). Renewal was also indicated by bilateral insula and dACC activation (CS+>CS-) in the early renewal phases as well as diminished activation (CS->CS+) in bilateral amygdala, hippocampus, vmPFC and right dACC in the novel context and left hippocampus in the acquisition context.

For post hoc ratings for the three different contexts we determined the level of significance at α=.05 (two-tailed) with Bonferroni-Holm correction for multiple testing. They

indicated higher fear (main effect context Greenhouse-Geisser $F_{(1.7, 164.8)}$=5.59, p=.004), as well as lower valence (main effect context Greenhouse-Geisser $F_{(1.8, 180.7)}$=5.64, p=.004) but no difference in arousal (main effect context Greenhouse-Geisser $F_{(1.8, 174.7)}$=3.55, p=.03), for the acquisition context compared to the novel context (fear: $t_{(100)}$=3.45, p<.001; arousal: $t_{(101)}$=1.54, p=.126; valence: $t_{(101)}$=-3.26, p=.002), and also higher fear ratings compared to the extinction context (fear: $t_{(101)}$=2.99, p=.003; arousal: $t_{(100)}$=2.44, p=.016; valence: $t_{(101)}$=-2.69, p=.008; differences for arousal and valence did not reach the level of significance after Bonferroni-Holm correction). The extinction context and the novel context did not differ significantly from each other (all p>.18). Regarding UCS expectancy ratings, a main effect for context was observed (Greenhouse-Geisser $F_{(1.6, 161.6)}$=.4.60, p=.011). Here, UCS expectancy in the acquisition context was significantly higher than in the novel context ($t_{(101)}$=3.41, p>.001), but not compared to the extinction context ($t_{(101)}$=1.89, p=.062). Extinction and novel contexts did also not differ ($t_{(101)}$=.59, p=.557).

G) Additional analyses with BDI-II as covariate

Because the patients with clinically relevant intrusion symptoms (INT group) and the patients without these symptoms (NO-INT group) differed in their scores on the Beck Depression Inventory II (BDI-II; Hautzinger, Keller, & Kühner, 2006), we performed additional analyses on the significant activation differences between these groups again with this score as a covariate (see Supplementary Table 5). Stronger activation decrease in the INT group in the right amygdala during extinction training, as well as activation differences in the left vmPFC during extinction recall in the extinction context (INT<NO-INT) and during renewal in the novel context C (INT<NO-INT) are still present. This suggests that the activation differences are not due to different scores on the BDI-II. On the other hand the differences in neural activation in the left vmPFC during the late phase of extinction (INT<NO-INT) is no longer significant and in the right amygdala during renewal in the acquisition context (INT>NO-INT) only tends to be significant (p=.070). Accordingly, the latter two effects could also be driven by the different scores of the BDI-II in the two groups. The groups did not differ in the frequency of depression diagnoses, which were assessed via the Diagnostic Interview for Mental Disorders (Margraf,

Cwik, Suppiger, & Schneider, 2017), which was administered by two trained psychologists. Nevertheless, the INT group in self-report indicates suffering more from depression symptoms, which could have an influence on the contextual modulation of fear influence. The literature on fear conditioning in patients with depression is sparse. In particular, according to our knowledge, there are no studies on the contextual modulation of conditioned fear and recall/renewal in different contexts. The literature on fear extinction processes in depression is mixed. Some studies find enhanced extinction (Kuhn et al., 2014), some studies find impaired extinction (Dibbets, van den Broek, & Evers, 2015; Wurst et al., 2021) or little support for either direction (Adolph, Teismann, Wannemüller, & Margraf, 2023). In the discussion for impaired extinction are cognitive deficits (Lee, Hermens, Porter, & Redoblado-Hodge, 2012). Future studies could consider cognitive abilities as a covariate in their analyses. On the other hand, enhancement of synaptic plasticity in the amygdala in patients with depression is discussed as a possible reason for improved extinction (Kuhn et al., 2014).

Supplementary Tables

Table 3. *Exclusion criteria for patients with social anxiety disorder and healthy controls*

Instruments	Criterion
A) SAD and HC	
Questionnaires and interview	age <18 or >65
	psychology student >second semester
	serious physical or neurological illness
	current psychotherapeutic treatment
	use of psychotropic drugs or drugs affecting the central nervous system in the past 12 weeks
	consumption of illegal drugs (current or regular in the past)
	alcohol consumption four or more times per week for a duration of one year
	MRI exclusion criteria for scientific measures
	color blindness
DIPS	acute suicidal tendencies
LEC-5	has experienced a traumatic event in the last four weeks
CAPS-5	meets criteria for PTSD
B) SAD	
LSAS, SPIN, DIPS, SKID-II	does not meet criteria for social anxiety disorder or has a LSAS sum score of ≤30 and a SPIN sum score of <25
	meets criteria for a primary disorder other than social anxiety disorder
	meets criteria for schizophrenia spectrum or other psychotic disorder, depressive episode/bipolar disorder with psychotic features, severe depressive episode, acute stress disorder, or moderate or severe substance use disorder
	meets criteria for emotionally unstable personality disorder
C) HC	
Questionnaire and interview	past psychotherapeutic treatment
LSAS, SPIN, DIPS, SKID-II	has a LSAS sum score of >30 and a SPIN sum score of ≥25
	meets or met criteria for any mental disorder

Note. SAD=social anxiety disorder, HC=healthy controls; LSAS=Liebowitz Social Anxiety Scale, SPIN=Social Phobia, DIPS (Diagnostic Interview for Mental Disorders), SKID-II (Structured clinical interview for DSM-IV Axis II: Personality Disorders), LEC-5 (life-event checklist for DSM-5), CAPS-5=Clinician-administered PTSD scale.

Table 4. *Neural activation differences for CS+ minus CS- and CS- minus CS+ during fear acquisition in context A (A), during extinction learning (B) and during the second half of extinction (C) in context B, during early extinction recall in context B (D) and during the second half of extinction (C) in context B, during early extinction recall in context B (D) and during early fear renewal in context C (E) and context A (F)*

Brain structure	*x*	*y*	*z*	T_{max}	p_{corr}
A) Fear acquisition					
CS+ minus CS-					
R amygdala	22	-2	-12	5.934	<.001
L amygdala	-20	-2	-14	5.215	<.001
R insula	34	20	-8	6.470	<.001
L insula	-30	22	8	8.714	<.001
R dACC	2	6	44	8.688	<.001
L dACC	-6	10	42	10.458	<.001
R vmPFC	18	14	-12	5.667	<.001
L vmPFC	-14	16	-12	5.984	<.001
R hippocampus	24	-22	-12	3.728	.025
L hippocampus	-20	-20	-16	4.501	.002
	-26	-40	-2	3.514	.045
CS- minus CS+					
L vmPFC	-2	34	-24	4.038	.018
B) Extinction learning					
Decrease in CS+ minus CS- (first half minus second half)					
L vmPFC	-18	12	-14	3.824	.037
Increase in CS+ minus CS- (second half minus first half)					
no significant results					
C) Second half of extinction learning					
CS+ minus CS-					
L insula	-30	26	0	4.104	.013
CS- minus CS+					
R amygdala	22	-12	-12	3.405	.032
L amygdala	-22	-8	-18	3.544	.020
R insula	36	-14	18	4.046	.015
R vmPFC	2	48	-12	4.923	<.001
L vmPFC	0	48	-14	4.606	.002
D) Early extinction recall					
CS+ minus CS-					
R insula	32	24	4	7.292	<.001
L insula	-30	24	6	8.945	<.001
R dACC	2	14	32	6.296	<.001

L dACC	-8	16	34	6.625	<.001
CS- minus CS+					
R amygdala	22	-8	-18	4.860	<.001
L amygdala	-30	-6	-20	4.399	.001
L vmPFC	0	36	-26	4.354	.007
R hippocampus	22	-10	-18	4.776	<.001
L hippocampus	-22	-12	-18	4.647	.001
E) Early fear renewal in the novel context					
CS+ minus CS-					
R insula	32	24	4	6.741	<.001
L insula	-30	26	4	6.836	<.001
R dACC	6	-16	30	4.784	.002
	6	32	22	4.219	.011
L dACC	-6	6	40	4.957	<.001
CS- minus CS+					
R amygdala	26	-8	-18	4.265	.002
L amygdala	-20	-8	-16	4.479	.001
R dACC	12	-14	48	3.780	.042
R vmPFC	6	56	-12	4.010	.022
L vmPFC	-6	58	-12	4.486	.005
R hippocampus	26	-10	-18	4.583	.001
L hippocampus	-24	-14	-18	4.013	.011
F) Early fear renewal in the acquisition context					
CS+ minus CS-					
R insula	30	20	-8	4.943	<.001
L insula	-30	18	10	5.215	<.001
R dACC	10	22	34	4.346	.007
	6	-16	30	4.291	.008
L dACC	-6	8	42	3.791	.034
CS- minus CS+					
L hippocampus	-32	-32	-12	4.466	.002

Note. L=left, R=right. The significance threshold was set to $p \leq .05$ (FWE-corrected). All coordinates (*x, y, z*) are given in MNI space. The probability masks for amygdala, insula and hippocampus are taken from the current 'Harvard-Oxford Cortical and Subcortical Structural Atlases` (Harvard Center for Morphometric Analysis [http://www.cma.mgh.harvard.edu/], probability threshold: 0.50) included in the FSL software package (http://www.fmrib.ox.ac.uk/fsl/). Masks for dACC and vmPFC were constructed with the MARINA software package (Walter, 2002).

Table 5. *Neural activation differences (CS+ minus CS-) for significant results from the main analyses between groups of patients with intrusion symptoms and patients without intrusion symptoms (see Figure 3) using depression symptoms (as measured by the BDI-II) as a covariate during extinction learning (A) and during the second half of extinction (B) in context B, during early extinction recall in context B (C) and during early fear renewal in context C (D) and context A (E)*

Brain structure	*x*	*y*	*z*	T_{max}	p_{corr}
A) Extinction learning: Decrease in CS+ minus CS- (first minus second half)					
INT>NO-INT					
R amygdala	26	-4	-22	3.49	.040
B) Second half of extinction learning					
INT<NO-INT					
L vmPFC	-10	48	-18	3.33	.173
C) Early extinction recall					
INT<NO-INT					
L vmPFC	-8	22	-24	4.01	.037
D) Early fear renewal in the novel context					
INT<NO-INT					
L vmPFC	-8	38	-8	4.22	.024
E) Early fear renewal in the acquisition context					
INT>NO-INT					
R amygdala	14	-8	-14	3.81	.070

Note. L=left, R=right. The significance threshold was set to $p \leq .05$ (FWE-corrected). All coordinates (*x, y, z*) are given in MNI space. The probability masks for amygdala are taken from the current 'Harvard-Oxford Cortical and Subcortical Structural Atlases` (Harvard Center for Morphometric Analysis [http://www.cma.mgh.harvard.edu/], probability threshold: 0.50) included in the FSL software package (http://www.fmrib.ox.ac.uk/fsl/). Masks for vmPFC were constructed with the MARINA software package (Walter, 2002).

Supplementary Figures

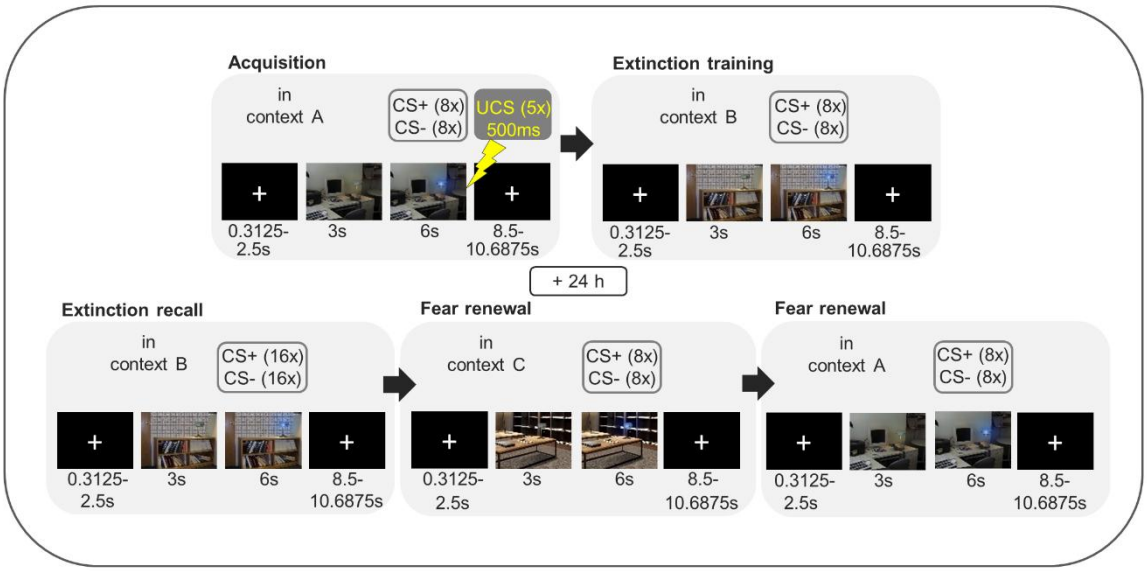

Figure 4. Schematic illustration of the fear conditioning paradigm.

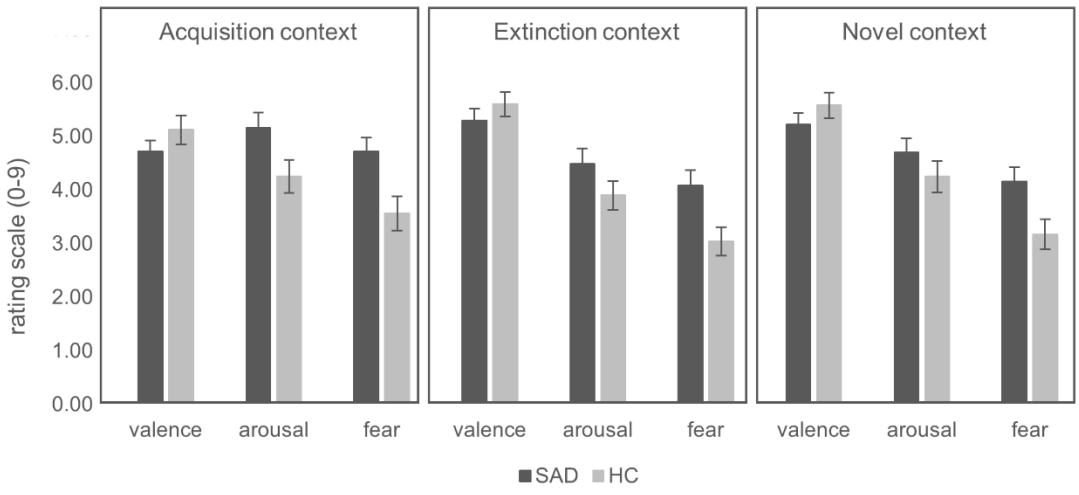

Figure 5. Post hoc ratings of valence, arousal and fear for the different contexts in patients with social anxiety disorder (SAD) and healthy controls (HC). Error bars represent standard errors of the mean.

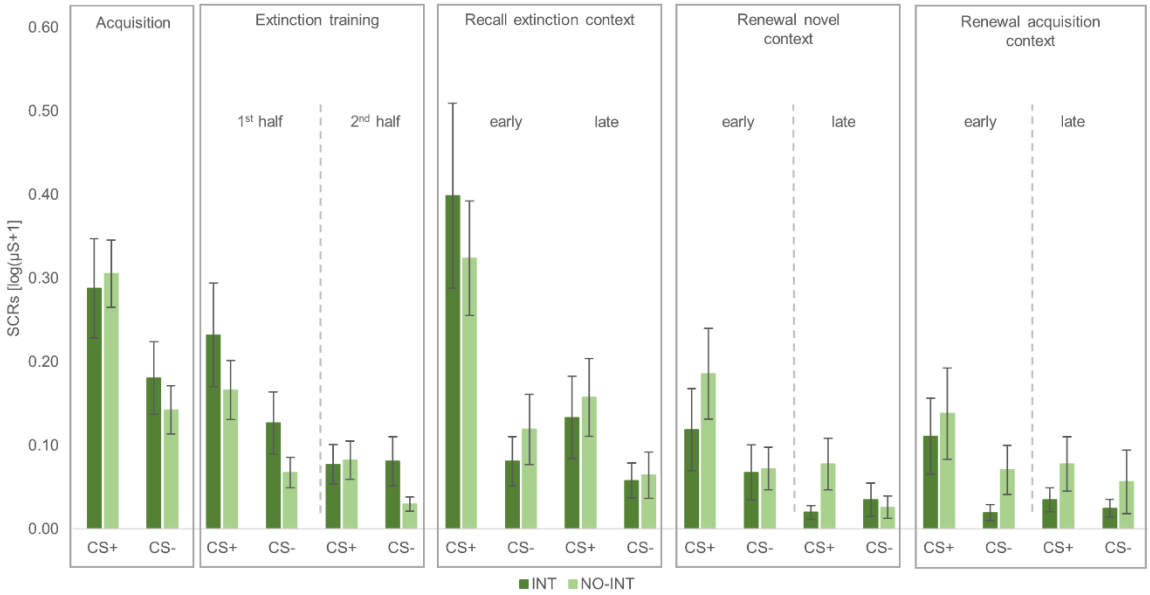

Figure 6. Skin conductance responses for CS+ and CS- during all phases of fear conditioning in social anxiety disorder patients with clinically relevant intrusions and patients without intrusions in response to an aversive social event. INT=patients with clinically relevant intrusions; NO-INT=patients without intrusions; µS=Microsiemens. '1st half' refers to the first eight trials, '2nd half' to the last eight trials of both CSs. The last two trials of each CS of the extinction training, which were used for the analyses of spontaneous recovery in the extinction context, are not illustrated here. 'Early' refers to the first four trials, 'late' to the last two trials of both CSs. Error bars represent standard errors of the mean.

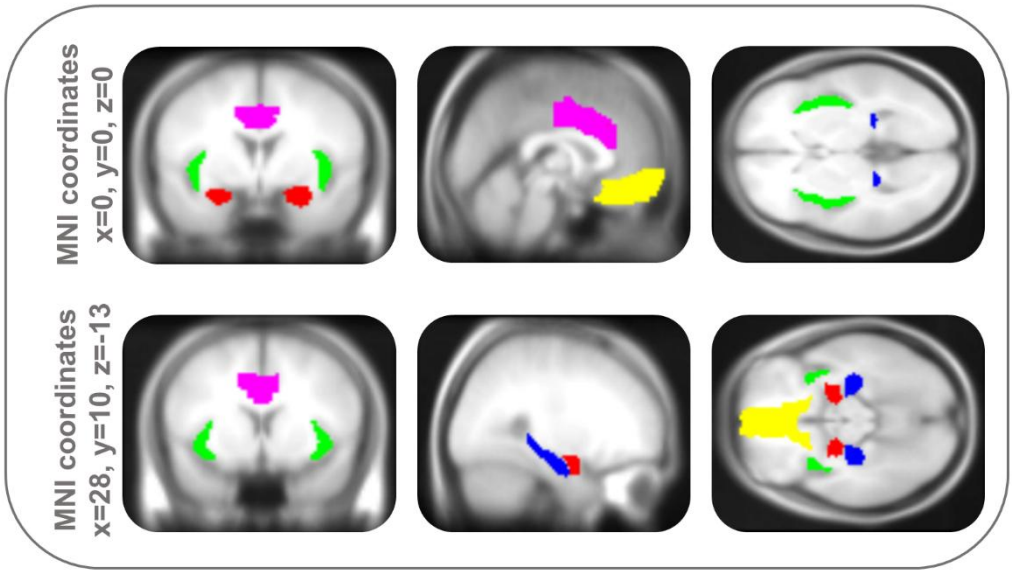

Figure 7. Combined mask with all regions of interest (dACC=violet; insula=green; vmPFC=yellow; amygdala=red; hippocampus=blue). Coordinates (x, y, z) are given in MNI space.

References

Adolph, D., Teismann, T., Wannemüller, A., & Margraf, J. (2023). Fear conditioning and extinction learning in the mood and anxiety disorders spectrum - Associations with the outcome of cognitive behavior therapy. *Behaviour Research and Therapy*, *160*, 104229. https://doi.org/10.1016/j.brat.2022.104229

Dibbets, P., van den Broek, A., & Evers, E. A. T. (2015). Fear conditioning and extinction in anxiety- and depression-prone persons. *Memory (Hove, England)*, *23*(3), 350–364. https://doi.org/10.1080/09658211.2014.886704

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Das Beck Depressions Inventar II. Deutsche Bearbeitung und Handbuch zum BDI-II [Beck Depression Inventory II. German edition and handbook for BDI-II]*. Harcourt Test Services.

Hermann, A., Stark, R., Müller, E. A., Kruse, O., Wolf, O. T., & Merz, C. J. (2020). Multiple extinction contexts modulate the neural correlates of context-dependent extinction learning and retrieval. *Neurobiology of Learning and Memory*, *168*, 107150. https://doi.org/10.1016/j.nlm.2019.107150

Kuhn, M., Höger, N., Feige, B., Blechert, J., Normann, C., & Nissen, C. (2014). Fear extinction as a model for synaptic plasticity in major depressive disorder. *PLOS ONE*, *9*(12), e115280. https://doi.org/10.1371/journal.pone.0115280

Lee, R. S. C., Hermens, D. F., Porter, M. A., & Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. *Journal of Affective Disorders*, *140*(2), 113–124. https://doi.org/10.1016/j.jad.2011.10.023

Margraf, J., Cwik, J. C., Suppiger, A., & Schneider, S. (2017). *DIPS Open Access 1.2: Diagnostisches Interview bei psychischen Störungen*. Ruhr-Universität Bochum (RUB). https://doi.org/10.13154/RUB.100.89

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology*, *46*(5), 984–995. https://doi.org/10.1111/j.1469-8986.2009.00852.x

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Seinsche, R. J., Fricke, S., Schäfer, A., Neudert, M. K., Zehtner, R. I., Stark, R., & Hermann, A. (2023). Effects of Imagery Rescripting on Emotional Responses During Imagination of a Socially Aversive Experience. *Journal of Emotion and Psychopathology*, *1*(1), 113–128. https://doi.org/10.55913/joep.v1i1.12

Sperl, M. F. J., Panitz, C., Rosso, I. M., Dillon, D. G., Kumar, P., Hermann, A., . . . Mueller, E. M. (2019). Fear Extinction Recall Modulates Human Frontomedial Theta and Amygdala Activity. *Cerebral Cortex*, *29*(2), 701–715. https://doi.org/10.1093/cercor/bhx353

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Walter, B. (2002). MARINA : An easy to use tool for the creation of MAsks for Region of INterest Analyses. *The 9th International Conference on Functional Mapping of the Human Brain, Available on CD-Rom in NeuroImage*, *19*(2).

Wurst, C., Schiele, M. A., Stonawski, S., Weiß, C., Nitschke, F., Hommers, L., . . . Menke, A. (2021). Impaired fear learning and extinction, but not generalization, in anxious and non-anxious depression. *Journal of Psychiatric Research*, *135*, 294–301. https://doi.org/10.1016/j.jpsychires.2021.01.034

Studie 2: When neutral faces become negative: The role of self-referential processing in social anxiety disorder – an fMRI study [2]

Abstract

Background: The processing of faces appears to be altered in patients with social anxiety disorder (SAD), which might be related to enhanced self-referential processing (SRP) of social stimuli as e.g., faces. The present study aims to investigate the relevance of SRP for the processing of neutral faces and the associated neural correlates in SAD.

Methods: During a neutral face processing task, 57 patients with SAD and 59 healthy controls (HC) were instructed to watch neutral or scrambled neutral faces and to increase or decrease SRP of neutral faces during functional magnetic resonance imaging. Main outcome measures were blood oxygen level-dependent responses in regions of interest and valence ratings.

Results: The SAD group reported reduced valence while watching neutral compared to scrambled faces in comparison to HC, but no significant differences in neural activation were observed between groups for this comparison. Increasing SRP led to decreased valence and stronger activation in insula and anterior cingulate cortex in patients with SAD compared to HC, whereas decreasing SRP elicited increased dlPFC activation without affecting valence ratings.

Conclusions: Our study provides evidence that patients with SAD are characterized by a stronger negative affective response towards neutral faces which are augmented during increased SRP. This is complemented by altered neural correlates during the processing of neutral faces under increased or decreased SRP compared to HC. These findings underscore the importance of therapeutic approaches to improve facial expression recognition and to reduce SRP.

Introduction

Current models of social anxiety disorder (SAD) emphasize the importance of self-focused attention (Clark & Wells, 1995) and attention towards external social cues as crucial

[2] Das Manuskript zu Studie 2 wird zur Publikation vorbereitet.

sources of information for the evaluation of one's own performance in social situations (Rapee & Heimberg, 1997). The process of being aware that a specific content is related to oneself is called self-referential processing (SRP; for example 'Tom looks grim, he's probably mad at me') (Northoff, 2011; Northoff et al., 2011). SRP and other cognitive processes, as self-focused attention, which occur in the context of social situations in SAD, can lead to the maintenance of social phobic symptoms, as they can result in more negative thoughts and anxiety, reduce the likelihood of discovering positive external cues or have a negative impact on performance (Burgio et al., 1986; Norton & Abbott, 2016; Spurr & Stopa, 2002; Woody & Rodriguez, 2000). Consequently, changing processes such as self-focused attention (Schreiber et al., 2015) and SRP e.g., via cognitive restructuring (Rodebaugh et al., 2004; for example 'Tom looks grim, he´s probably having a bad day') are important interventions in the treatment of SAD.

At the neural level, self-referential processes are found to be reflected in activity in cortical midline structures (CMS: orbitomedial prefrontal cortex, dorsomedial prefrontal cortex, anterior cingulate cortex [ACC], precuneus/posterior cingulate cortex [PCC]) (Northoff et al., 2006), where the medial prefrontal cortex (Araujo et al., 2013; Gusnard et al., 2001; Moran et al., 2009; Whitfield-Gabrieli et al., 2011) and especially the ventromedial prefrontal cortex (vmPFC) seem to be of particular importance (D. D. Wagner et al., 2019; Yaoi et al., 2015). A recent meta-analysis (H.-J. Yoon et al., 2019) summarized functional magnet resonance imaging (fMRI) studies on SRP in SAD in which participants had to read or assign self-related or other-related comments or statements (Blair et al., 2008; Goldin & Gross, 2010; Goldin et al., 2009; H.-J. Yoon et al., 2016), read 1st or 2nd viewpoint comments (Blair et al., 2011), anticipate a public and evaluated speech versus reading a word (Boehme et al., 2014), read and empathize in social situations with intentional/unintentional social transgressions (Blair et al., 2010), perform a visual search task while being exposed to internal or external threat (Choi et al., 2016), or look at faces with/without eye contact (Schneier et al., 2011) to manipulate SRP. Patients with SAD compared to HC showed hyperactivation of CMS, temporal pole (TP), temporo-parietal junction (TPJ) and insula during SRP, emphasizing the role of these brain regions for altered SRP in SAD.

As described in the explanatory model of Rapee and Heimberg (1997), one focus of attention in patients with SAD is on external cues which could possibly be an indicator of negative evaluation, such as signs of boredom in the other person. Important external stimuli that patients with SAD tend to interpret in a self-referential manner and as threatening are other people's faces. The processing of faces is therefore the subject of numerous studies in the context of social anxiety (Gilboa-Schechtman & Shachar-Lavie, 2013; Machado-de-Sousa et al., 2010), whereby neutral faces are often used as a baseline condition for emotional facial expressions in these studies. However, previous findings show that neutral facial expressions are typically not perceived as neutral (Lewinski, 2015). This is especially the case for patients with mental disorders in general (Filkowski & Haas, 2017), as well as for socially anxious individuals (Peschard & Philippot, 2017; K. L. Yoon & Zinbarg, 2008) and patients with SAD in particular (Lange et al., 2012). However, there are just a few studies investigating the neural correlates of neutral face processing compared to other (baseline) conditions, like a fixation condition (Cooney et al., 2006), scrambled faces (Gentili et al., 2008) or an aversive odor (Birbaumer et al., 1998), showing reduced activation in the left fusiform gyrus (Gentili et al., 2008) or stronger activation in the amygdala (Birbaumer et al., 1998; Cooney et al., 2006) in patients with SAD compared to healthy controls. Even though these analyses focused on the comparison between neutral faces and a control condition, the results are confounded with the concurrent presentation of emotional faces or other aversive stimuli within the same experimental paradigm.

However, it is unknown if experimentally induced enhanced SRP is associated with a more negative perception of neutral faces in SAD. Therefore, the aim of this functional magnetic resonance imaging (fMRI) study in 57 patients with SAD and 59 HC was to examine the neural correlates of neutral face processing (without presentation of potentially confounding emotional faces) and the influence of SRP in a newly developed neutral face processing task. Participants were instructed to increase and decrease SRP of neutral facial expressions. Neutral face processing (watching neutral compared to scrambled faces) should be related to reduced valence (more unpleasant) and higher activation of amygdala, mPFC,

ACC, PCC, precuneus, TPJ and insula in SAD patients compared with HC. Increasing SRP is expected to lead to reduced valence and stronger activation of brain areas associated with self-referential processing (mPFC, ACC, PCC, precuneus, TPJ and insula) specifically in patients with SAD compared to HC. Moreover, we assume altered neural activation (in the abovementioned regions of interest) and valence changes while decreasing SRP in patients with SAD compared to HC without specifying a direction of the effect.

Methods and Materials

Subjects

We recruited participants with SAD and healthy controls (HC) via a mailing list of the local university, the associated outpatient clinic, flyers, and announcements in newspapers and on websites. Participants were invited to a first assessment of eligibility where, among others, the in- and exclusion criteria were checked. Patients should have a primary diagnosis of SAD; HC should not fulfil criteria for diagnosis of any current or past mental disorders (for detailed criteria see Supplementary Table 7). Throughout the project we had a drop-out of 15 patients with SAD and 8 HC (cancellation without reason [6 SAD/2 HC], scheduling reasons [3 SAD/3 HC], discomfort in scanner [5 SAD/2 HC], technical difficulties [1 SAD/1 HC]). Our final samples of 57 patients and 59 HC are comparable in age, sex, and number of years of education (see all relevant demographic and clinical variables for both groups in Table 6). This study is part of a larger project investigating the association of intrusive reexperiencing of aversive social events and the neural correlates of context-dependent fear conditioning in SAD (Fricke et al., 2024), as well as the effect of Imagery Rescripting of aversive memories (Seinsche et al., 2024).

Two trained psychologists administered the Diagnostic Interview for Mental Disorders (DIPS; Margraf et al., 2017) in order to reliably assign participants to the group of patients with SAD or HC. In addition, participants were interviewed using the Liebowitz Social Anxiety Scale (LSAS; Consbruch et al., 2016; Fresco et al., 2001; Liebowitz, 1987), which was also used as an inclusion criterion for the patient sample (SAD: total score≥30) or as an exclusion criterion

for the healthy sample (HC: total score<30 for inclusion). Additionally, we used the Social Phobia Inventory (SPIN; Connor et al., 2000; Consbruch et al., 2016; Sosic et al., 2008) where the total score had to be ≥25 to verify classification as a SAD patient and <25 as a HC.

The local ethics approved the study protocol, which is in line with the Declaration of Helsinki. All participants gave written informed consent and received ten euros/hour or course credit as compensation.

Table 6. *Demographic and clinical variables for patients with social anxiety disorder and healthy controls*

variable	SAD *n*=57	HC *n*=59	test for group differences
age, *M*(*SD*)	26.96(5.99)	27.24(7.41)	
range (years)	18-48	20-57	t=-.217(114), p=.828
sex, female/male	37/20	38/21	χ^2(1)=.003, p=.955
years of education, *M*(*SD*)	17.90(3.17)	17.74(3.10)	t=2.9(114), p=.776
symptoms of SAD,			
LSAS *M*(*SD*)	64.96(20.53)	10.55(7.52)	t=18.47(68.0), p<.001
SPIN *M*(*SD*)	38.54(9.40)	7.37(5.44)	t=21.77(89.1), p<.001
current comorbid diagnoses (without personality disorders), %			
none	50.9	100	
one	35.1	0	
two or more	14.0	0	
current comorbid personality disorder diagnoses, %			
none	70.2	100	
one	28.1	0	
two	1.8	0	
psychotherapeutic treatment in the past, %	35.1	0	
psychotropic drugs in the past, %	10.5	0	

Note. Table 6 shows descriptive variables separated for groups (SAD=social anxiety disorder, HC=healthy controls) and test statistics for group differences. *T*-test, degrees of freedom (*df*), χ^2-statistics, significance level (*p*), LSAS=Liebowitz Social Anxiety Scale, SPIN=Social Phobia Inventory, diagnoses were assessed by using the DIPS (Diagnostic Interview for Mental Disorders) and the SKID-II (Structured clinical interview for DSM-IV Axis II: Personality Disorders), sample sizes varied for LSAS: SAD *n*=55, HC *n*=56 and BDI-II: HC *n*=56.

Procedure

The above-mentioned larger project comprised a total of six appointments (for an overview of the entire project see Supplement A). Diagnostic and other interviews were conducted at the first appointment, which served, among other things, to categorize subjects into the two groups. At the fourth appointment, which took place within the following 15 days, the fMRI paradigm regarding processing of neutral faces was performed.

Stimuli

The facial stimuli were obtained from the Radboud Faces Database (Langner et al., 2010). We selected neutral faces from Caucasian adult models with neutral emotional expressions, eyes directed straight ahead and frontal camera position. As we needed only 36 (half female/half male) out of 39 appropriate models, we excluded 3 models based on deviating characteristics (red face colour, open mouth) and gender. The pictures were tailored so that they ended up in a 680x680 pixel format. As comparison condition we used phase-scrambled images of these faces, so that they have the same pixels but without facial content. For a training session we needed nine more faces that were selected randomly (five female/four male) from the Karolinska Data Set (Lundqvist et al., 1998; Oosterhof & Todorov, 2008) and had neutral emotional expressions, a direct gaze and were taken from a frontal position as well.

Paradigm

The experimental paradigm was presented using the Presentation® software (Neurobehavioral Systems, Inc.). At the beginning of each trial, we presented a white fixation cross on a black background jittered between 625ms and 2500ms. Afterwards, an instruction keyword showed up for 2000ms. There were three different instructions in German language that introduced either to ’watch attentively’ (’aufmerksam betrachten’), to imagine that the person, they were going to see ’thinks about me’ (’denkt über mich nach’) or ’thinks of something else’ (’denkt an etwas anderes’). With the last two instructions we intended to manipulate SRP (increase of SRP: ’thinks about me’ and decrease of SRP: ’thinks of

something else'). These latter two conditions were always followed by faces, never by scrambled pictures. Faces or scrambled pictures were presented each for 4000ms in a block of three pictures. Afterwards participants were asked to directly rate their valence with a keypad on a nine-point Likert scale (self-assessment manikin; SAM; Bradley & Lang, 1994) from unpleasant to pleasant. They had a 5000ms time interval to complete the rating. If they did not press a confirmation button within this time frame, the last recorded cursor position was used as the answer. At the end of each trial, a fixation cross was presented with the duration jittered between 1000 and 2875ms, up to a total block length of 22,500ms. The entire procedure of 13 minutes consisted of 32 blocks, within which each condition was presented eight times, as well as 30s of fixation cross presentation at the beginning and end of the experiment. The 36 faces were a priori divided into three sets à 12 pictures and the assignment of picture sets to conditions was counterbalanced across participants and groups.

Directly after the experimental task (outside the scanner) participants should indicate how successful they could put the three different instructions into practice (9-point scale from 'not at all' to 'very well'), respectively. Furthermore, we asked the participants to indicate valence, arousal (both SAM) and fear on 9-point scales ('unpleasant' to 'pleasant'/ 'calm' to 'excited'/ 'not at all' to 'very strong') for every set of 12 faces, in order to capture the lasting effects, the different instructions had on the emotional experience regarding the neutral faces. Participants were instructed to look at every single face of the set before answering these questions.

Functional Magnetic Resonance Imaging

Using a 3 Tesla whole body scanner (Siemens Prisma, Erlangen, Germany) and a 64 channel standard head coil, the BOLD responses were recorded with a T2*-weighted gradient echo-planar imaging sequence (EPI). We measured 42 slices with a slice thickness of 3mm (0.75mm gap; descending slice order; TE=30ms; TR=2.5s; flip angle=81°; Field of View=220x220mm²; matrix-size=74x74 pixel; parallel imaging: GRAPPA acceleration 2; oblique slices -30 degrees tilt from the AC-PC line). The functional image acquisition consisted of 312 volumes, whereby the first four volumes were excluded from analyses due to an

incomplete steady state of magnetization. Prior to the functional run, a gradient echo field map sequence was measured to receive information for unwarping.

Statistical Parametric Mapping (SPM12, r7219, Wellcome Department of Cognitive Neurology, London, UK; 2009), implemented in Matlab R2019a (Mathworks Inc., Sherborn, MA, USA) was used for data preprocessing and analysis. The preprocessing of the data consisted of the following: realignment, unwarping, slice time correction, coregistration, segmentation (T1: MPRAGE, sagittal, 0.9mm x 0.9mm x 0.9mm), normalization of the brain (standard space of the Montreal Neurological Institute), and smoothing with an isotropic three-dimensional Gaussian filter with a full-width at half maximum of 6mm.

Statistical Analyses

Afterwards, first and second level analyses were conducted. In the first level model, we modeled following regressors of interest based on the canonical hemodynamic response function (HRF): instruction (duration: 2s), ratings (individual durations), the implementation phases of the four different instructions (’watch attentively’ to neutral faces (*neutral)*, ’watch attentively’ to scrambled faces (*scrambled*), ’thinks about me’ (*highSRP*), ’thinks of something else’ (*lowSRP*)) (duration: 12s). For filtering voxel-based time series, we used a high-pass filter of 128s. As regressors of no interest we used the six movement parameters from realignment and one regressor for each volume with a framewise displacement >0.5mm (Power et al., 2012).

We contrasted the condition *neutral* with the other conditions (*neutral* > *scrambled*; *highSRP* > *neutral*; *lowSRP* > *neutral*) on an individual level. To compare the two groups, two-sample *t*-tests were performed in second-level analyses for the abovementioned contrasts. Amygdala, ACC, posterior cingulate gyrus, precuneus, dorsolateral and dorsomedial prefrontal cortex (dlPFC, dmPFC), ventrolateral and ventromedial prefrontal cortex (vlPFC, vmPFC), TPJ and insula were defined as regions of interest (ROI; bilaterally in each case; probability masks for amygdala, ACC, precuneus and insula taken from the current ‘Harvard-Oxford Cortical and Subcortical Structural Atlases`, Harvard Center for Morphometric Analysis [http://www.cma.mgh.harvard.edu/], included in the FSL software package

(http://www.fmrib.ox.ac.uk/fsl/); masks for dlPFC, dmPFC, vlPFC and vmPFC were created with the MARINA software package (Walter, 2002), based on the AAL parcellation (Tzourio-Mazoyer et al., 2002), and the TPJ mask with the Brainnetome atlas (Fan et al., 2016). We used the small volume correction option of SPM12 and a significance threshold of α≤.05 on voxel level with family-wise error (FWE) correction and an intensity threshold of *p*=.05, uncorrected. The reported coordinates (x, y, z) refer to MNI space. Exploratory whole brain analyses were conducted with a significance threshold of *α*≤.05 on voxel-level corrected for multiple testing for the whole brain (FWE-correction). The minimal cluster size (*k*) for exploratory analyses was 5 voxels.

The online-rating data were analyzed with IBM SPSS 29 (IBM Corporation, Armonk, USA). We calculated mean values for each condition and conducted mixed-design analyses of variance (ANOVA) with *group* as between subject factor (SAD, control subjects) and *condition* as within subject factor (for the hypothesis regarding neutral faces processing: *scrambled*, *neutral*; for the hypothesis regarding SRP: *neutral* and *highSRP*/*lowSRP, respectively*). In the same way, six ANOVAs were carried out for the post-hoc ratings (success in implementing the instructions; lasting effects concerning valence, arousal and fear with two conditions each: *neutral* vs *highSRP* and *neutral* vs *lowSRP*). For post-hoc tests we determined the level of significance at *α*=.05 with Bonferroni-Holm correction for multiple testing.

Results

Group comparisons showed that patients with SAD and control subjects did not differ regarding age, gender and years of education, but in social anxiety severity (LSAS, SPIN; see Table 6).

Neutral Face Processing

For watching neutral faces compared to scrambled pictures, there were no significant activation differences between groups (*neutral vs. scrambled)*.

Valence ratings (see Figure 8) showed a significant interaction effect of *condition* and *group* (Greenhouse-Geisser $F_{(1,114)}$=5.69, p=.019) with the SAD group reporting reduced valence compared to the HC group in the *neutral* condition (SAD: M=5.08(0.78), HC: M=5.52(0.97), $t_{(114)}$=2,69, p=.004, one-tailed), but no difference between groups in the *scrambled* condition (SAD: M=5.36(1.50), HC: M=5.17(1.21), $t_{(114)}$=-0.76, p=.48, two-tailed).

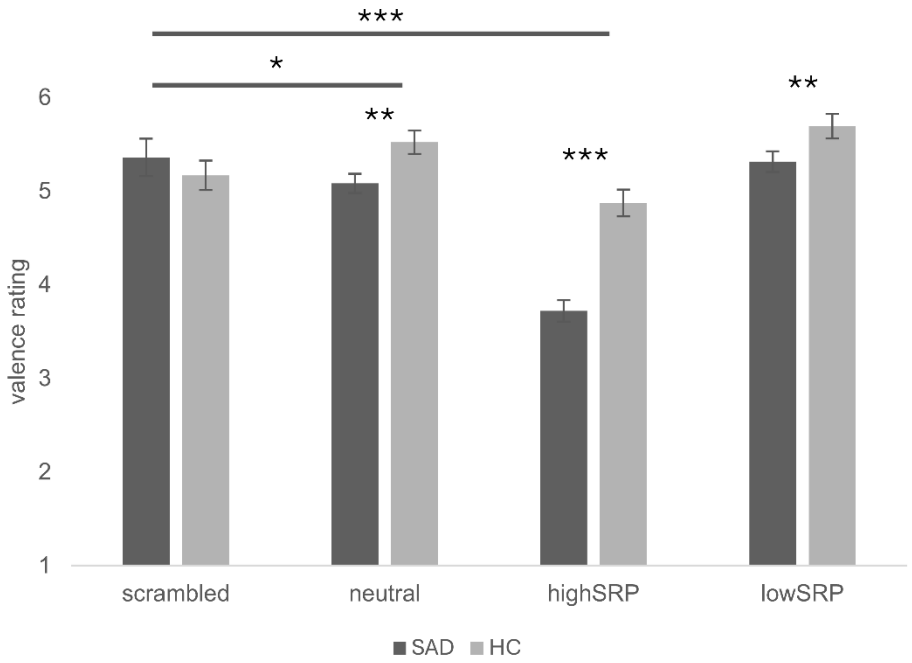

Figure 8. Valence ratings on a 9-point scale (’unpleasant’ to ’pleasant’) during the face processing task for the four different conditions. Neutral=’watch attentively’ to neutral faces, scrambled=’watch attentively’ to scrambled faces, highSRP=’thinks about me’ lowSRP=’thinks of something else’, SAD=social anxiety disorder, HC=healthy controls.

Increased self-referential processing

During increased self-referential processing (*highSRP>neutral*) the SAD group revealed stronger activation in the ACC (T_{max}=3.97, $p_{FWEcorr}$=.031, MNI: x=6, y=22, z=30) and insula (T_{max}=4.12, $p_{FWEcorr}$=.024, MNI: x=-32, y=10, z=-14; T_{max}=3.99, $p_{FWEcorr}$=.036, MNI: x=-30, y=14, z=-10; T_{max}=3.96, $p_{FWEcorr}$=.040, MNI: x=-42, y=2, z=-2; T_{max}=3.90, $p_{FWEcorr}$=.048, MNI: x=-36, y=6, z=-2) in comparison to HC, while no further activation differences could be detected, neither in ROI nor in exploratory whole brain analyses (see Figure 9).

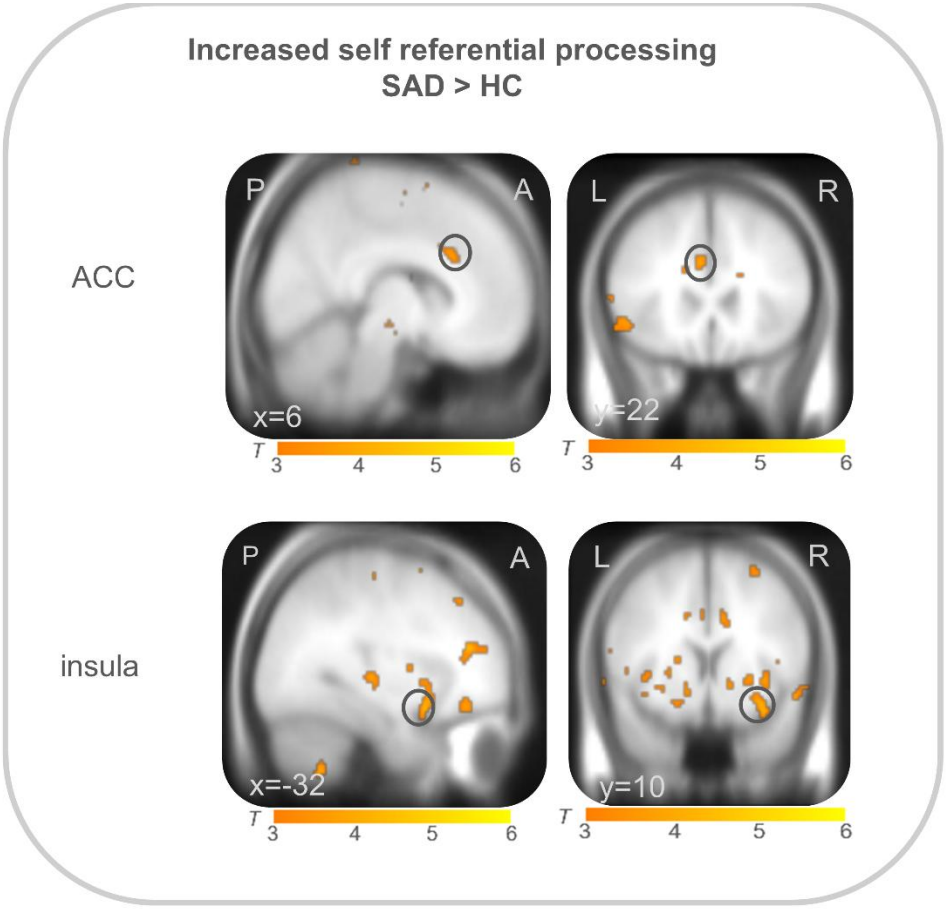

Figure 9. Stronger neural activation in the ACC and insula in patients with social anxiety disorder in comparison to healthy controls during increased self-referential processing (highSRP>neutral). Figure shows stronger neural activation in the ACC (T_{max}=3.97, $p_{FWEcorr}$=.031, MNI: x=6, y=22, z=30) and insula (T_{max}=4.12, $p_{FWEcorr}$=.024, MNI: x=-32, y=10, z=-14; T_{max}=3.99, $p_{FWEcorr}$=.036, MNI: x=-30, y=14, z=-10; T_{max}=3.96, $p_{FWEcorr}$=.040, MNI: x=-42, y=2, z=-2; T_{max}=3.90, $p_{FWEcorr}$=.048, MNI: x=-36, y=6, z=-2) in patients with social anxiety disorder in comparison to healthy controls. Coordinates are given in MNI space; color bars depict T-values; activations were superimposed on the MNI305 T1 template; SAD=social anxiety disorder, HC=healthy controls.

The ANOVA for valence ratings (*highSRP* vs *neutral)* showed significant main effects for *condition* ($F_{(1, 114)}$=111,54, *p*<.001) and *group* ($F_{(1, 114)}$=29.64, *p*<.001) and their interaction ($F_{(1, 114)}$=13,94, *p*<.001). Enhanced SRP compared to watching neutral leads to a stronger reduction of valence ratings in the SAD compared to the HC group (see Figure 8). Post-hoc tests revealed a significant reduced valence in both groups (SAD: *neutral: M*=5.08(0.78), *highSRP: M*=3.72(0.88), $t_{(56)}$=-10,64, *p*<.001; HC: *neutral: M*=5.52(0.97), *highSRP: M*=4.87(1.09), $t_{(58)}$=-4.62, *p*<.001, both two-tailed), while patients showed reduced valence during both, enhanced SRP and watching neutral compared to HC (*highSRP*: $t_{(114)}$=6.24, *p*<.001, *neutral*: $t_{(114)}$=2.69, *p*=.004; both one-tailed).

Decreased self-referential processing

The reduction of self-referential processing (*lowSRP>neutral*) led to increased activation in the dlPFC (T_{max}=4.51, $p_{FWEcorr}$=.026, MNI: x=38, y=52, z=14) in patients with SAD compared to HC, while all other ROI analyses for this contrast revealed no significant results (see Figure 10). Exploratory whole brain analyses for *lowSRP>neutral* moreover revealed stronger activation in the Heschl's gyrus (T_{max}=5.60, $p_{FWEcorr}$=.008, MNI: x=-50, y=-18, z=0) and central opercular cortex (T_{max}=5.17, $p_{FWEcorr}$=.042, MNI: x=-48, y=-12, z=10).

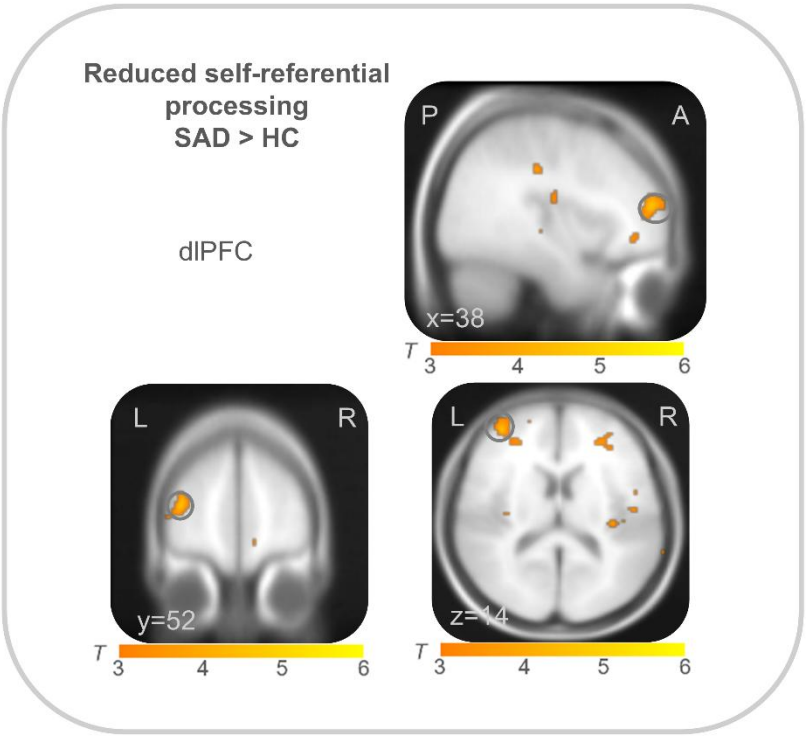

Figure 10. Stronger neural activation in the dlPFC in patients with social anxiety disorder in comparison to healthy controls during decreased self-referential processing (lowSRP>neutral). Figure shows stronger neural activation in the dlPFC (T_{max}=4.51, $p_{FWEcorr}$=.026, MNI: x=38, y=52, z=14) in patients with social anxiety disorder in comparison to healthy controls. Coordinates are given in MNI space; color bars depict T-values; activations were superimposed on the MNI305 T1 template; SAD=social anxiety disorder, HC=healthy controls.

The ANOVA for valence ratings (*neutral*, *lowSRP*) revealed significant main effects for *condition* ($F_{(1, 114)}$=6.82, p=.010) and *group* ($F_{(1, 114)}$=7.61, p=.007) with reduced valence in the *neutral* condition compared to the *lowSRP* condition ($t_{(115)}$=2.62, p=.010, two-tailed) and reduced valence in both conditions in the SAD group in comparison to HC (*neutral: SAD* M=5.08(0.78), *HC* M=5.52(0.97), $t_{(114)}$=2.69, p=.004; *lowSRP*: SAD M=5.32(0.83), HC

M=5.69(1.00), $t_{(113)}$=2.72, *p*=.004; both one-tailed), while there was no significant interaction effect ($F_{(1, 114)}$=.16, *p*=.70).

Lasting effects

The ANOVAs for the lasting effects (see Figure 11) concerning valence ratings revealed significant main effects of *group* (*neutral* vs *highSRP* $F_{(1, 113)}$=7.21, *p*=.008; *neutral* vs *lowSRP* $F_{(1, 113)}$=5.90, *p*=.017) with the patients with SAD reporting significant reduced valence in the two SRP manipulating conditions compared to HC (*highSRP* $t_{(114)}$=3.38, *p*<.001; *lowSRP* $t_{(113)}$=2.72, *p*=.004; *neutral* $t_{(113)}$=1.46, *p*=.073; all one-tailed). Analyses of lasting effects concerning arousal and fear ratings both showed significant main effects of group (arousal: *neutral* vs *highSRP* $F_{(1, 113)}$=19.54, *p*<.001; *neutral* vs *lowSRP* $F_{(1, 113)}$=20.10, *p*<.001; fear: *neutral* vs *highSRP* $F_{(1, 113)}$=25.54, *p<.001*; *neutral vs lowSRP* $F_{(1, 113)}$=18.70, *p*<.001) with the patient group reporting more arousal and fear in all conditions (arousal: *highSRP* $t_{(114)}$=-4.38, *p*<.001; *lowSRP* $t_{(113)}$=-4.46, *p*<.001; *neutral* $t_{(113)}$=-3.88, *p*<.001; fear: *highSRP* $t_{(114)}$=-5.00, *p*<.001; *lowSRP* $t_{(113)}$=-3.69, *p*<.001; *neutral* $t_{(113)}$=-4.29, *p*<.001; all one-tailed). There were no further significant main or interaction effects (valence: context main effect $F_{(2, 226)}$=.42, *p*=.66, interaction effect $F_{(2, 226)}$=1.25, *p*=.29; arousal: context main effect $F_{(2, 226)}$=.49, *p*=.62; interaction effect $F_{(2, 226)}$=.54, *p*=.59, fear: context main effect $F_{(2, 226)}$=.26, *p*=.77; interaction effect $F_{(2, 226)}$=.42, *p*=.66).

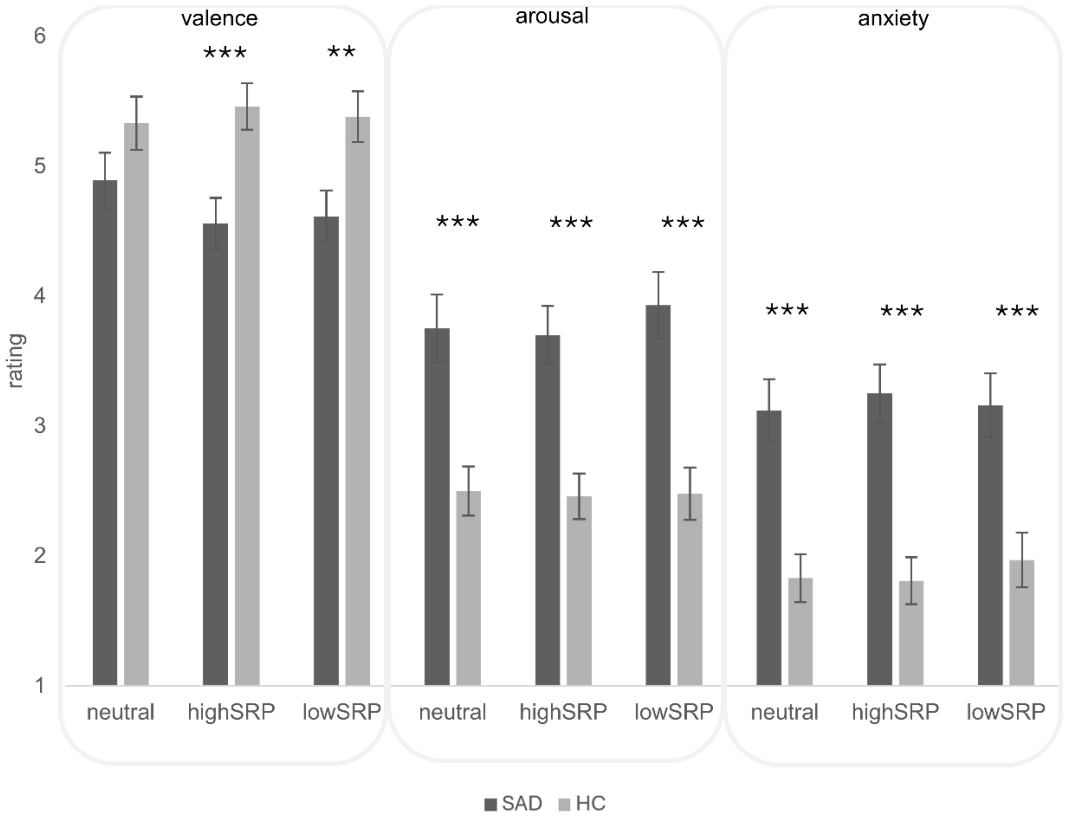

Figure 11. Valence, arousal and fear ratings after the face processing task. Figure shows lasting effects of the different instructions on the emotional experience of neutral faces on 9-point scales (valence: 'unpleasant' to 'pleasant', arousal: 'calm' to 'excited', fear: 'not at all' to 'very strong'). 'Watch attentively' with faces= neutral, 'thinks about me'=highSRP, 'thinks of something else'=lowSRP, SAD=social anxiety disorder, HC=healthy controls.

Success ratings

Overall, most of the subjects rated the implementation of the instructions as successful (responses on a 9-point scale from 'not at all' to 'very well'; SAD: *neutral M*=7.39(1.52), *highSRP M*=6.96(1.58), *lowSRP M*=6.82(1.57); HC: *neutral M*=7.92(1.12), *highSRP M*=7.49(1.41), *lowSRP M*=7.19(1.35)).

Discussion

To further our understanding of neutral face processing and the influence of SRP in SAD, we used a newly developed neutral face processing task without the presentation of potentially confounding emotional faces as mostly done in previous studies. While the SAD group reported reduced valence during the processing of neutral faces compared to HC, there were, however, no significant group differences at the neural level. In accordance with our

hypothesis, enhanced SRP led to stronger activation of ACC and insula and a stronger reduction of self-reported valence in patients with SAD compared to HC. During decreased SRP we found stronger dlPFC activation in patients with SAD vs HC but no differential effects on valence ratings.

Watching neutral faces was associated with reduced self-reported valence in patients with SAD compared to HC. This is consistent with our hypothesis and with previous findings in which subjects with social anxiety were more likely to identify neutral faces as negative (Peschard & Philippot, 2017; Winton et al., 1995) and perceive them as threatening (Lange et al., 2012). However, we could not demonstrate this difference at the neural level. Previous studies reported stronger activation in patients with SAD during neutral face processing in the amygdala (Birbaumer et al., 1998; Cooney et al., 2006) and the fusiform gyrus (Gentili et al., 2008). In contrast to our study, these studies also used emotional faces or aversive stimuli in their paradigms. As a result, the response to neutral faces in patients with SAD, primed by aversive facial expressions or stimuli, may have been stronger in these studies. Neural activation in patients with SAD in response to emotional faces has already been well studied, and differences in activation have been found, including the amygdala (Gentili et al., 2016) and the fusiform gyrus (Gentili et al., 2008) compared to HC. The missing activation difference for watching neutral faces in the current study might also be related to methodological issues (e.g., duration of stimulus presentation, block design). However, our results show that patients with SAD interpret neutral faces as more unpleasant compared to HC.

An explanatory model of SAD describes how these patients often perceive external stimuli, such as faces, in a self-referential manner and thus as a potentially negative evaluation of themselves (Rapee & Heimberg, 1997). To specifically investigate the influence of SRP during the processing of neutral faces, we furthermore manipulated the self-reference of neutral faces in our experimental paradigm. When SRP was increased, the valence decreased in both groups, which is consistent with previous findings (Mor & Winquist, 2002; Woody & Rodriguez, 2000). As expected, the effect of increased SRP on valence ratings was stronger in the patient group than in the HC group, which is in line with studies showing stronger effects

of SRP on affective responses in clinical samples (Mor & Winquist, 2002) and the abovementioned explanatory model of SAD (Rapee & Heimberg, 1997). As in previous studies, increased SRP was associated with stronger activation of the ACC and insula in social anxiety (Boehme et al., 2014; Choi et al., 2016; Goldin & Gross, 2010; Goldin et al., 2009; H.-J. Yoon et al., 2016). Furthermore, previous studies showed insula activation in association with interoceptive perception (Critchley et al., 2004) and associated social anxiety (Terasawa et al., 2013). Cognitive explanatory models also address the tendency toward increased interoceptive perception as a critical factor contributing to the maintenance of SAD (Rapee & Heimberg, 1997). The ACC, as part of the CMS, was associated with SRP (Northoff et al., 2006) and appeared to be hyperactivated during increased SRP in SAD (Choi et al., 2016; Goldin & Gross, 2010; Goldin et al., 2009), which was replicated by the findings of our study. In line with this, a further study (G. Wagner et al., 2013) demonstrated increased activation in the rostral ACC in healthy participants, particularly in response to negative self-related information. Our results could therefore indicate that the nature of self-related thoughts was more negative in patients with SAD compared with HC. In contrast to previous findings, our data showed no hyperactivity in other regions of the CMS, such as mPFC, PCC, precuneus and TPJ. One difference between our study and previous studies that found differences in these regions (e.g., Blair et al., 2008; Goldin & Gross, 2010; Goldin et al., 2009; H.-J. Yoon et al., 2016), is the applied paradigm. While in other studies the valence of self-reference was specified (e.g., positive or negative adjectives, praise or criticism, negative self-beliefs), we did not determine it. Therefore, variability in the type and valence of self-related thoughts/information may have led to these differing findings. In summary, our findings confirm the hypothesis that increased SRP has a stronger negative effect in patients with SAD compared to HC. Therefore, modifying increased SRP or its negative effect is an important goal in psychotherapy for SAD (Rodebaugh et al., 2004; Schreiber et al., 2015).

The reduction of SRP led to more positive valence in all participants than simply watching the faces, with patients in both conditions reporting lower valence than the HC. The valence rating thus did not confirm our hypothesis that the two groups would differ in their

response to reduced SRP. However, at the neural level, stronger activation in the dlPFC was observed in patients with SAD during reduced SRP. The (d)lPFC is central for cognitive control (Koechlin et al., 2003) and plays a role in cognitive bias (Xia et al., 2021). Lemogne et al. (2009) found stronger functional connectivity of the dlPFC with the dACC and medial frontal gyrus during SRP ('self') compared to a control condition ('general') in patients with depression. They suggested that the strong self-focus of these patients may have required further cognitive control by the dlPFC for emotion regulation (Ochsner & Gross, 2005). In subjects with higher (social) anxiety, stronger activation of the dlPFC was also associated with emotional control (Bramson et al., 2023; Zhuang et al., 2025). Our own results may suggest that although patients with SAD were able to regulate their emotions during our paradigm by reducing self-reference, they seem to have to invest more effort to do so, which may be reflected in stronger activation of the dlPFC. However, they still did not reach the level of HC, as they reported reduced valence compared to HC.

The selection of our paradigm represents both a strength and limitation of our study. Although our success ratings ensured that participants were able to implement the instructions, the concrete underlying process remains unclear (i.e., thoughts and actions). A follow-up interview would have been useful to obtain this information. At the same time, our manipulation of self-reference through the open instruction may be particularly representative of the individual processes involved in confronting neutral faces in everyday life (increased SRP) or in implementing SRP reduction as a therapeutic strategy in patients.

Our study shows how severely patients with SAD are already affected by neutral or ambivalent social stimuli as they report reduced valence during the processing of neutral faces compared to HC. The analysis of lasting effects also shows that patients experience greater negative affective responses when watching neutral faces in the aftermath. These results underscore the importance of therapeutic approaches that aim to improve the ability to recognize facial expressions. Enhanced self-referential processing had a stronger negative effect on valence in patients with SAD. Moreover, reducing self-reference probably required more effort for patients (stronger activation of dlPFC) and still resulted in lower valence than in

HC. Accordingly, our results also highlight the importance of therapeutic approaches that aim to reduce SRP, such as perspective taking, cognitive reality checks, or behavioral experiments to check dysfunctional cognitions. Our finding of hyperactive insula in patients with SAD during increased SRP, which is also associated with interoceptive perception (Critchley et al., 2004; Terasawa et al., 2013), further supports the relevance of training to redirect attention (e.g., from interoceptive perception to the current demand). In future studies on the processing of neutral faces, the selection of paradigms and comparison conditions should take into account that the processing of neutral faces could be confounded by other potentially aversive stimuli, such as emotional faces.

References

Araujo, H. F., Kaplan, J., & Damasio, A. (2013). Cortical midline structures and autobiographical-self processes: An activation-likelihood estimation meta-analysis. *Frontiers in Human Neuroscience*, *7*, 548. https://doi.org/10.3389/fnhum.2013.00548

Birbaumer, N., Grodd, W., Diedrich, O., Klose, U., Erb, M., Lotze, M., Schneider, F., Weiss, U., & Flor, H. (1998). fMRI reveals amygdala activation to human faces in social phobics. *NeuroReport*, *9*(6).

Blair, K. S., Geraci, M., Devido, J., McCaffrey, D., Chen, G., Vythilingam, M., Ng, P., Hollon, N., Jones, M., & Pine, D. S. (2008). Neural response to self- and other referential praise and criticism in generalized social phobia. *Archives of General Psychiatry*, *65*(10), 1176–1184. https://doi.org/10.1001/archpsyc.65.10.1176

Blair, K. S., Geraci, M., Hollon, N., Otero, M., Devido, J., Majestic, C., Jacobs, M., Blair, K. S., & Pine, D. S. (2010). Social norm processing in adult social phobia: Atypically increased ventromedial frontal cortex responsiveness to unintentional (embarrassing) transgressions. *The American Journal of Psychiatry*, *167*(12), 1526–1532. https://doi.org/10.1176/appi.ajp.2010.09121797

Blair, K. S., Geraci, M., Otero, M., Majestic, C., Odenheimer, S., Jacobs, M., & Pine, D. S. (2011). Atypical modulation of medial prefrontal cortex to self-referential comments in generalized social phobia. *Psychiatry Research*, *193*(1), 38–45. https://doi.org/10.1016/j.pscychresns.2010.12.016

Boehme, S., Ritter, V., Tefikow, S., Stangier, U., Strauss, B., Miltner, W. H. R., & Straube, T. (2014). Brain activation during anticipatory anxiety in social anxiety disorder. *Social Cognitive and Affective Neuroscience*, *9*(9), 1413–1418. https://doi.org/10.1093/scan/nst129

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

Bramson, B., Meijer, S., van Nuland, A., Toni, I., & Roelofs, K. (2023). Anxious individuals shift emotion control from lateral frontal pole to dorsolateral prefrontal cortex. *Nature Communications*, *14*(1), 4880. https://doi.org/10.1038/s41467-023-40666-3

Burgio, K. L., Merluzzi, T. V., & Pryor, J. B. (1986). Effects of performance expectancy and self-focused attention on social interaction. *Journal of Personality and Social Psychology*, *50*(6), 1216–1221. https://doi.org/10.1037//0022-3514.50.6.1216

Choi, S.-H., Shin, J.-E., Ku, J., & Kim, J.-J. (2016). Looking at the self in front of others: Neural correlates of attentional bias in social anxiety. *Journal of Psychiatric Research*, *75*, 31–40. https://doi.org/10.1016/j.jpsychires.2016.01.001

Clark, D. M., & Wells, A. (1995). A Cognitive Model of Social Phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social Phobia: Diagnosis, Assessment, and Treatment* (pp. 69–93). THE GUILFORD PRESS. https://books.google.de/books?hl=de&lr=&id=rXrekuSy2bsC&oi=fnd&pg=PA69&dq=clark+wells+social+phobia+model&ots=gSdK2kuUcs&sig=fLuK2DiSmxKJIm3ZiK1A1qGDbIU#v=onepage&q=clark%20wells%20social%20phobia%20model&f=false

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Foa, E. B., & Weisler, R. H. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *The British Journal of Psychiatry: The Journal of Mental Science*, *176*, 379–386. https://doi.org/10.1192/bjp.176.4.379

Consbruch, K., Stangier, U., & Heidenreich, T. (2016). *Soziale Angst-Skalen (SOZAS). SIAS, SPS, SPIN, LSAS*. Hogrefe.

Cooney, R. E., Atlas, L. Y., Joormann, J., Eugène, F., & Gotlib, I. H. (2006). Amygdala activation in the processing of neutral faces in social anxiety disorder: Is neutral really neutral? *Psychiatry Research*, *148*(1), 55–59. https://doi.org/10.1016/j.pscychresns.2006.05.003

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A [Arne], & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195. https://doi.org/10.1038/nn1176

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex (New York, N.Y. : 1991)*, *26*(8), 3508–3526. https://doi.org/10.1093/cercor/bhw157

Filkowski, M. M., & Haas, B. W. (2017). Rethinking the Use of Neutral Faces as a Baseline in fMRI Neuroimaging Studies of Axis-I Psychiatric Disorders. *Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging*, *27*(3), 281–291. https://doi.org/10.1111/jon.12403

Fresco, D. M., Coles, M. E., Heimberg, R. G., Liebowitz, M. R., Hami, S., Stein, M. B., & Goetz, D. (2001). The Liebowitz Social Anxiety Scale: A comparison of the psychometric properties of self-report and clinician-administered formats. *Psychological Medicine*, *31*(6), 1025–1035. https://doi.org/10.1017/S0033291701004056

Fricke, S., Seinsche, R. J., Neudert, M. K., Schäfer, A., Zehtner, R. I., Stark, R., & Hermann, A. (2024). Neural correlates of context-dependent extinction recall in social anxiety disorder: Relevance of

intrusions in response to aversive social experiences. *Psychological Medicine*, *54*(3), 548–557. https://doi.org/10.1017/S0033291723002179

Gentili, C., Cristea, I. A., Angstadt, M., Klumpp, H., Tozzi, L., Phan, K. L., & Pietrini, P. (2016). Beyond emotions: A meta-analysis of neural response within face processing system in social anxiety. *Experimental Biology and Medicine (Maywood, N.J.)*, *241*(3), 225–237. https://doi.org/10.1177/1535370215603514

Gentili, C., Gobbini, M. I., Ricciardi, E., Vanello, N., Pietrini, P., Haxby, J. V., & Guazzelli, M. (2008). Differential modulation of neural activity throughout the distributed neural system for face perception in patients with Social Phobia and healthy subjects. *Brain Research Bulletin*, *77*(5), 286–292. https://doi.org/10.1016/j.brainresbull.2008.08.003

Gilboa-Schechtman, E., & Shachar-Lavie, I. (2013). More than a face: A unified theoretical perspective on nonverbal social cue processing in social anxiety. *Frontiers in Human Neuroscience*, *7*, 904. https://doi.org/10.3389/fnhum.2013.00904

Goldin, P. R., & Gross, J. J. (2010). Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. *Emotion (Washington, D.C.)*, *10*(1), 83–91. https://doi.org/10.1037/a0018441

Goldin, P. R., Ramel, W., & Gross, J. J. (2009). Mindfulness Meditation Training and Self-Referential Processing in Social Anxiety Disorder: Behavioral and Neural Effects. *Journal of Cognitive Psychotherapy*, *23*(3), 242–257. https://doi.org/10.1891/0889-8391.23.3.242

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4259–4264. https://doi.org/10.1073/pnas.071043098

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science (New York, N.Y.)*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Lange, W.-G., Allart, E., Keijsers, G. P. J., Rinck, M., & Becker, E. S. (2012). A neutral face is not neutral even if you have not seen it: Social anxiety disorder and affective priming with facial expressions. *Cognitive Behaviour Therapy*, *41*(2), 108–118. https://doi.org/10.1080/16506073.2012.666563

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

Lemogne, C., Le Bastard, G., Mayberg, H., Volle, E., Bergouignan, L., Lehéricy, S., Allilaire, J.-F., & Fossati, P. (2009). In search of the depressive self: Extended medial prefrontal network during self-referential processing in major depression. *Social Cognitive and Affective Neuroscience*, *4*(3), 305–312. https://doi.org/10.1093/scan/nsp008

Lewinski, P. (2015). Automated facial coding software outperforms people in recognizing neutral faces as neutral from standardized datasets. *Frontiers in Psychology*, *6*, 1386. https://doi.org/10.3389/fpsyg.2015.01386

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, *22*, 141–173. https://doi.org/10.1159/000414022

Lundqvist, D., Flykt, A., & Ohman, A [A.]. (1998). *Karolinska Directed Emotional Faces: Database of standardized facial images.* Karolinska Hospital.

Machado-de-Sousa, J. P., Arrais, K. C., Alves, N. T., Chagas, M. H. N., Meneses-Gaya, C. de, Crippa, J. A. d. S., & Hallak, J. E. C. (2010). Facial affect processing in social anxiety: Tasks and stimuli. *Journal of Neuroscience Methods*, *193*(1), 1–6. https://doi.org/10.1016/j.jneumeth.2010.08.013

Margraf, J., Cwik, J. C., Suppiger, A., & Schneider, S. (2017). *DIPS Open Access 1.2: Diagnostisches Interview bei psychischen Störungen.* Universitätsbibliothek Bochum. https://doi.org/10.13154/RUB.100.89

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, *128*(4), 638–662. https://doi.org/10.1037//0033-2909.128.4.638

Moran, J. M., Heatherton, T. F., & Kelley, W. M. (2009). Modulation of cortical midline structures by implicit and explicit self-relevance evaluation. *Social Neuroscience*, *4*(3), 197–211. https://doi.org/10.1080/17470910802250519

Neurobehavioral Systems, Inc. *Presentation* (Version 19.0) [Computer software]. www.neurobs.com

Northoff, G. (2011). Self and brain: What is self-related processing? *Trends in Cognitive Sciences*, *15*(5), 186-7; author reply 187-8. https://doi.org/10.1016/j.tics.2011.03.001

Northoff, G., Heinzel, A., Greck, M. de, Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. https://doi.org/10.1016/j.neuroimage.2005.12.002

Northoff, G., Qin, P., & Feinberg, T. E. (2011). Brain imaging of the self--conceptual, anatomical and methodological issues. *Consciousness and Cognition*, *20*(1), 52–63. https://doi.org/10.1016/j.concog.2010.09.011

Norton, A. R., & Abbott, M. J. (2016). Self-Focused Cognition in Social Anxiety: A Review of the Theoretical and Empirical Literature. *Behaviour Change*, *33*(1), 44–64. https://doi.org/10.1017/bec.2016.2

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Peschard, V., & Philippot, P. (2017). Overestimation of threat from neutral faces and voices in social anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, *57*, 206–211. https://doi.org/10.1016/j.jbtep.2017.06.003

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, *35*(8), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3

Rodebaugh, T. L., Holaway, R. M., & Heimberg, R. G. (2004). The treatment of social anxiety disorder. *Clinical Psychology Review*, *24*(7), 883–908. https://doi.org/10.1016/j.cpr.2004.07.007

Schneier, F. R., Pomplun, M., Sy, M., & Hirsch, J. (2011). Neural response to eye contact and paroxetine treatment in generalized social anxiety disorder. *Psychiatry Research*, *194*(3), 271–278. https://doi.org/10.1016/j.pscychresns.2011.08.006.

Schreiber, F., Heimlich, C., Schweitzer, C., & Stangier, U. (2015). Cognitive therapy for social anxiety disorder: The impact of the "self-focused attention and safety behaviours experiment" on the course of treatment. *Behavioural and Cognitive Psychotherapy*, *43*(2), 158–166. https://doi.org/10.1017/S1352465813000672

Seinsche, R. J., Fricke, S., Neudert, M. K., Zimmer, R. I., Stark, R., & Hermann, A. (2024). Emotional Changes during Imagery Rescripting of Aversive Social Memories in Social Anxiety Disorder: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *93*(4), 264–270. https://doi.org/10.1159/000539402

Sosic, Z., Gieler, U., & Stangier, U. (2008). Screening for social phobia in medical in- and outpatients with the German version of the Social Phobia Inventory (SPIN). *Journal of Anxiety Disorders*, *22*(5), 849–859. https://doi.org/10.1016/j.janxdis.2007.08.011

Spurr, J. M., & Stopa, L. (2002). Self-focused attention in social phobia and social anxiety. *Clinical Psychology Review*, *22*(7), 947–975. https://doi.org/10.1016/S0272-7358(02)00107-1

Terasawa, Y., Shibata, M., Moriguchi, Y., & Umeda, S. (2013). Anterior insular cortex mediates bodily sensibility and social anxiety. *Social Cognitive and Affective Neuroscience*, *8*(3), 259–266. https://doi.org/10.1093/scan/nss108

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Wagner, D. D., Chavez, R. S., & Broom, T. W. (2019). Decoding the neural representation of self and person knowledge with multivariate pattern analysis and data-driven approaches. *Wiley Interdisciplinary Reviews. Cognitive Science*, *10*(1), e1482. https://doi.org/10.1002/wcs.1482

Wagner, G., Koch, K., Schachtzabel, C., Peikert, G., Schultz, C. C., Reichenbach, J. R., Sauer, H., & Schlösser, R. G. (2013). Self-referential processing influences functional activation during cognitive control: An fMRI study. *Social Cognitive and Affective Neuroscience*, *8*(7), 828–837. https://doi.org/10.1093/scan/nss074

Walter, B. (2002). MARINA : An easy to use tool for the creation of MAsks for Region of INterest Analyses. *The 9th International Conference on Functional Mapping of the Human Brain, Available on CD-Rom in NeuroImage*, *19*(2).

Whitfield-Gabrieli, S., Moran, J. M., Nieto-Castañón, A., Triantafyllou, C., Saxe, R., & Gabrieli, J. D. E. (2011). Associations and dissociations between default and self-reference networks in the human brain. *NeuroImage*, *55*(1), 225–232. https://doi.org/10.1016/j.neuroimage.2010.11.048

Winton, E. C., Clark, D. M., & Edelmann, R. J. (1995). Social anxiety, fear of negative evaluation and the detection of negative emotion in others. *Behaviour Research and Therapy*, *33*(2), 193–196. https://doi.org/10.1016/0005-7967(94)E0019-F

Woody, S. R., & Rodriguez, B. F. (2000). Self-Focused Attention and Social Anxiety in Social Phobics and Normal Controls. *Cognitive Therapy and Research*, *24*(4), 473–488. https://doi.org/10.1023/A:1005583820758

Xia, X., Li, Y., Wang, Y., Xia, J., Lin, Y., Zhang, X., Liu, Y., & Zhang, J. (2021). Functional role of dorsolateral prefrontal cortex in the modulation of cognitive bias. *Psychophysiology*, *58*(10), e13894. https://doi.org/10.1111/psyp.13894

Yaoi, K., Osaka, M., & Osaka, N. (2015). Neural correlates of the self-reference effect: Evidence from evaluation and recognition processes. *Frontiers in Human Neuroscience*, *9*, 383. https://doi.org/10.3389/fnhum.2015.00383

Yoon, H.-J., Kim, J. S., Shin, Y.-B., Choi, S.-H., Lee, S.-K., & Kim, J.-J. (2016). Neural activity during self-referential working memory and the underlying role of the amygdala in social anxiety disorder. *Neuroscience Letters*, *627*, 139–147. https://doi.org/10.1016/j.neulet.2016.05.068

Yoon, H.-J., Seo, E. H., Kim, J.-J., & Choo, I. H. (2019). Neural Correlates of Self-referential Processing and Their Clinical Implications in Social Anxiety Disorder. *Clinical Psychopharmacology and Neuroscience : The Official Scientific Journal of the Korean College of Neuropsychopharmacology*, *17*(1), 12–24. https://doi.org/10.9758/cpn.2019.17.1.12

Yoon, K. L., & Zinbarg, R. E. (2008). Interpreting neutral faces as threatening is a default mode for socially anxious individuals. *Journal of Abnormal Psychology*, *117*(3), 680–685. https://doi.org/10.1037/0021-843X.117.3.680

Zhuang, Q., Yao, S., Xu, L., Chen, S., Li, J., Zheng, X., Fu, M., Kendrick, K. M., & Becker, B. (2025). A functional anatomical shift from the lateral frontal pole to dorsolateral prefrontal cortex in emotion action control underpins elevated levels of anxiety – partial replication and generalization of Bramson et al., 2023. Advance online publication. https://doi.org/10.1101/2025.03.09.642129

Supplemental Information

A) Course of the entire project

The whole project consisted of six appointments. Diagnostic interviews and questionnaires, as well as interviews on various autobiographical events took place on a first appointment. Neuropsychological testing and an experiment on pattern separation were conducted on a second appointment. The third and fourth sessions, at which an fMRI paradigm on fear conditioning was performed, took place within 14 days of the diagnostic interviews and followed each other for approximately 24 hours. At the second scanning session, after the fear conditioning paradigm, the experiment described here on the processing of neutral faces

followed. The patient sample was invited for two more appointments to examine both, memory features and posttraumatic stress symptoms related to the autobiographical events, as well as the effect of Imagery Rescripting.

Supplementary Tables

Table 7. *Exclusion criteria for patients with social anxiety disorder and healthy controls*

Instruments	Criterion
A) **SAD and HC**	
Questionnaires and interview	age < 18 or > 65
	psychology student > second semester
	serious physical or neurological illness
	current psychotherapeutic treatment
	use of psychotropic drugs or drugs affecting the central nervous system in the past 12 weeks
	consumption of illegal drugs (current or regular in the past)
	alcohol consumption four or more times per week for a duration of one year
	MRI exclusion criteria for scientific measures
	color blindness
DIPS	acute suicidal tendencies
LEC-5	has experienced a traumatic event in the last four weeks
CAPS-5	meets criteria for PTSD
B) **SAD**	
LSAS, SPIN, DIPS, SKID-II	does not meet criteria for social anxiety disorder or has a LSAS sum score of ≤ 30 and a SPIN sum score of < 25
	meets criteria for a primary disorder other than social anxiety disorder
	meets criteria for schizophrenia spectrum or other psychotic disorder, depressive episode/bipolar disorder with psychotic features, severe depressive episode, acute stress disorder or moderate or severe substance use disorder
	meets criteria for emotionally unstable personality disorder
C) **HC**	
Questionnaire and interview	past psychotherapeutic treatment
LSAS, SPIN	has a LSAS sum score of > 30 and a SPIN sum score of ≥ 25
DIPS, SKID-II	meets or met criteria for any mental disorder

Note. SAD=social anxiety disorder, HC=healthy controls; LSAS=Liebowitz Social Anxiety Scale, SPIN=Social Phobia, DIPS (Diagnostic Interview for Mental Disorders), SKID-II (Structured clinical interview for DSM-IV Axis II: Personality Disorders), LEC-5 (life-event checklist for DSM-5), CAPS-5=Clinician-administered PTSD scale

Zusammenfassende Diskussion und Ausblick

Die vorliegende Dissertation widmet sich der Untersuchung verschiedener Prozesse der Bedrohungsverarbeitung bei SAD. Im Mittelpunkt stehen dabei zwei spezifische, bisher wenig oder ungenügend beleuchtete Aspekte. Zum einen die kontextabhängige Modulation konditionierter Furcht bei Patient:innen mit SAD (Fricke et al., 2024). Dabei wurden in Anlehnung an Befunde, welche bereits aus der Forschung zu PTBS vorliegen, Patient:innen mit intrusivem Wiedererleben in Reaktion auf ein auslösendes sozial aversives Ereignis mit Patient:innen ohne intrusives Wiedererleben verglichen. Zum anderen stand die Verarbeitung neutraler Gesichter, als Beispiel uneindeutiger sozialer Reize, unter dem Einfluss der Manipulation selbstbezogener Verarbeitung im Mittelpunkt. Hervorzuheben ist, dass für diese Dissertationsstudie ein neu entwickeltes Paradigma genutzt wurde, bei dem nur neutrale Gesichter als Reize dienten. In bisherigen Studien wurden meist auch emotionale Gesichter, oder aversive Reize eingesetzt, welche die Verarbeitung der neutralen Gesichter verfälscht haben könnten (Birbaumer et al., 1998; Cooney et al., 2006; Gentili et al., 2008). Damit liefert die Dissertation einen Mehrwert zur bisherigen Literatur, indem sie ein differenziertes Verständnis der neuronalen und kognitiven Mechanismen liefert, welche der Entstehung und Aufrechterhaltung der SAD zugrunde liegen könnten.

Die Ergebnisse der Dissertationsstudie vermittelten ein heterogenes Bild über die kontextabhängige Modulation konditionierter Furcht in SAD, von dem keine eindeutige Parallele zu PTBS in der Gesamtstichprobe der Patient:innen mit SAD gezogen werden kann. In der vorliegenden Stichprobe zeigte sich, dass Intrusionssymptome auch bei Patient:innen mit SAD relevante Symptome darstellen und häufig einen Bezug zu ätiologisch relevanten aversiven Ereignissen aufweisen. Bisher finden diese Symptome in ätiologischen Modellen von SAD hauptsächlich in der Art von verzerrten Vorstellungsbildern von sich selbst Beachtung (Heimberg et al., 2014), welche mit früheren Erlebnissen der Patient:innen assoziiert sein können (Norton & Abbott, 2017). Ätiologische Modelle der PTBS beschreiben Intrusionen als Folge unzureichender Kontextualisierung der traumatischen Erfahrung (Ehlers & Clark, 2000; Liberzon & Abelson, 2016), wofür es in der vorliegenden Stichprobe, insbesondere bei den

Patient:innen mit klinisch relevanten Intrusionssymptomen, ebenfalls Anhaltspunkte gab. Demnach deuten die Befunde darauf hin, dass Intrusionen bei SAD einen klinisch relevanten Subtyp kennzeichnen könnten. Daraus abgeleitet, wäre ein Vergleich der SAD-Untergruppen in ihrer Verarbeitung neutraler Gesichter und den Effekten selbstbezogener Verarbeitung interessant. Dieses Vorgehen ließe sich auch theoretisch begründen, da Intrusionen als primär selbstorientiert beschrieben werden (D. A. Clark & Rhyno, 2005) und demnach selbst als eine Form selbstbezogener Aufmerksamkeit konzeptualisiert werden könnten. In der vorliegenden Studie zur Verarbeitung neutraler Gesichter konnte gezeigt werden, dass eine Erhöhung des Selbstbezugs mit einer Reduktion der Valenz der Proband:innen beim Betrachten neutraler Gesichter einherging. Zusätzlich hängt selbstbezogene Aufmerksamkeit bei sozialer Angst mit negativem Affekt zusammen (Mor & Winquist, 2002). Zusammengenommen könnte daher davon ausgegangen werden, dass die negative Interpretation neutraler Gesichter in der Untergruppe der Patient:innen mit klinisch relevanten Intrusionen noch stärker vorhanden ist als bei Patient:innen ohne Intrusionen. Auch könnte diesen Patient:innen, deren Aufmerksamkeitsfokus häufiger selbstbezogen ist (im Sinne von klinisch relevanten Intrusionen), die Regulation des Selbstbezugs schwerer fallen.

Darüber hinaus wäre denkbar, dass sich Defizite in den beiden untersuchten Bereichen, kontextabhängige Modulation konditionierter Furcht und selbstbezogene Verarbeitung neutraler Gesichter, wechselseitig beeinflussen. Wenn Patient:innen mit SAD soziale Situationen aufgrund von Defiziten bei der Nutzung von Sicherheitssignalen bzw. Defiziten in der kontextabhängigen Modulation ihrer Furcht (Fricke et al., 2024) häufiger als unsicher oder bedrohlich wahrnehmen, könnte dies mit einer ständigen Erhöhung des Selbstbezugs einhergehen. Gemäß den kognitiven Modellen zur SAD (D. M. Clark & Wells, 1995; Heimberg et al., 2014), richten die Betroffenen in potentiell bedrohlichen Situationen ihre Aufmerksamkeit nach innen (Selbstbezug), um ihre physiologische Reaktion und ihr Verhalten, z. B. im Sinne eines Sicherheitsverhaltens, zu überprüfen. Diese verstärkte selbstbezogene Verarbeitung wiederum erschwert die Verarbeitung von externen Sicherheitssignalen oder allgemein Kontextfaktoren und dementsprechend die funktionale Modulation von Furcht. Bei

dieser Interaktion könnte der Mechanismus der Furchtgeneralisierung relevant sein, welcher bei Patient:innen mit Angststörungen dysfunktional ausgeprägt ist (Cooper et al., 2022). In Studie 1 zeigte sich dies bei der Konditionierung von Furcht in der Gesamtgruppe der Patient:innen in Form einer stärkeren Amygdala Reaktion auf den Sicherheitsreiz und in der Untergruppe mit klinisch relevanten Intrusionen in Bezug auf die unzureichende Nutzung der Kontextinformationen zur Modulation der Furcht. In Studie 2 zeigte sich in der Gruppe der Patient:innen im Vergleich zu HC eine negativere Valenz bei der Betrachtung neutraler Gesichter, was eine Generalisierung von Furcht auf eigentlich harmlose soziale Reize darstellen könnte. Um dieser Furchtgeneralisierung entgegenzuwirken könnten in der Therapie sozialer Ängste z. B. Diskriminationstrainings zum Einsatz kommen. Dieses führte in einer Studie mit gesunden Teilnehmer:innen zur Reduktion von Furchtgeneralisierung (Ginat-Frolich et al., 2017). Auch Salzano et al. (2023) empfehlen in ihrem Psychotherapiemodell das Training neuropsychologischer Funktionen. Wenngleich die Hypothese zur Interaktion von Defiziten in den Bereichen der kontextabhängigen Modulation von Furcht und selbstbezogener Verarbeitung einen vielversprechenden Ansatz zum Verständnis der Aufrechterhaltung sozialer Ängste liefert, bedarf sie zunächst weiterer empirischer Überprüfung.

Obwohl die vorliegenden Befunde und daraus abgeleiteten Implikationen fruchtbar sind, müssen sie vor dem Hintergrund methodischer Limitationen interpretiert werden. Beide Studien weisen Einschränkungen auf, die zum einen die verwendeten Paradigmen betreffen. Beim kontextabhängigen Furchtkonditionierungsparadigma wurde unmittelbar nach der Furchtakquisition das Extinktionstraining durchgeführt. Dies kann die Rückkehr konditionierter Furcht bei Präsentation des sicheren Kontextes (Extinktionsabruf) zur Folge haben (Merz et al., 2016), wie es sich bei der Analyse der Gesamtstichprobe tatsächlich zeigte. Das neu entwickelte Paradigma zur Verarbeitung neutraler Gesichter stellt einerseits ein besonders naturalistisches Design dar, da es Raum für individuelle Verarbeitungsprozesse, ohne Vorgabe konkreter gedanklicher Strategien, lässt. Andererseits bleiben daher die konkreten Vorgänge, welche während des Paradigmas ablaufen, ungewiss. Einen Anhaltspunkt bietet die Nachbefragung zum selbsteingeschätzten Erfolg bezüglich der Umsetzung der

Instruktionen. Dennoch hätte eine detailliertere Nachbefragung der Proband:innen mehr Aufschluss, beispielsweise über die konkreten Gedanken und Vorgehensweisen sowie eventuelle Unterschiede zwischen den Gruppen geben können. Auch das *block design* des Gesichter-Paradigmas könnte ein limitierender Faktor für die Entdeckung von Gruppenunterschieden gewesen sein (Chee et al., 2003). Beim Vergleich von Patient:innen mit und ohne Intrusionen zeigte sich ein Unterschied bezüglich depressiver Symptome, welche im Beck-Depressions-Inventar (BDI-II) angegeben wurden. Obwohl Patient:innen mit Intrusionen dort mehr dieser Symptome angaben, unterschieden sich die Gruppen nicht bezüglich des tatsächlichen Anteils von Patient:innen mit ausgeprägter Depression. Dennoch könnten auch solche depressiven Symptome, wie Konzentrationsschwierigkeiten, Rumination oder Selbstzweifel auf die Fähigkeit zur kontextabhängigen Modulation von Furcht oder die Wirkung selbstbezogener Verarbeitung Einfluss nehmen und damit Gruppenunterschiede bewirken. Zusätzliche Analysen mit den BDI-II Werten als Kovariate führten in der vorliegenden Studie allerdings überwiegend zu den gleichen Gruppenunterschieden. Dies unterstützt die Annahme, dass die Unterschiede in den beiden Subgruppen von Patient:innen mit SAD eher spezifisch auf die Heterogenität bezüglich Intrusionen zurückzuführen sind.

Trotz dieser methodischen Einschränkungen lassen sich aus der Dissertation über die spezifischen Befunde hinaus Rückschlüsse im größeren wissenschaftlichen Kontext ziehen. Die Heterogenität in der vorliegenden Stichprobe von Patient:innen mit SAD unterstreicht die Nützlichkeit transdiagnostischer Ansätze (z. B. Kotov et al., 2017). Intrusionssymptome, welche traditionell der PTBS zugeordnet werden (American Psychiatric Association, 2022), zeigen sich auch bei Patient:innen mit SAD in Reaktion auf sozial aversive Ereignisse und hängen mit gemeinsamen neuronalen Korrelaten (z. B. Hypoaktivierung vmPFC, Fricke et al., 2024; Marin et al., 2020; Milad et al., 2009; Rougemont-Bücking et al., 2011) zusammen. Bei SAD finden sich verzerrte Vorstellungsbilder von sich selbst in Erklärungsmodellen wieder (Ehlers & Clark, 2000; Heimberg et al., 2014), welche dem intrusiven Wiedererleben bei PTBS ähneln (Bjornsson et al., 2020; Ann Hackmann et al., 1998; Moscovitch et al., 2018). Als Mechanismus wird dabei bei PTBS eine unzureichende Kontextualisierung der traumatischen

Erinnerung beschrieben, welche zu kontinuierlicher Reaktivierung der Symptomatik führt (Ehlers & Clark, 2000). Die Ergebnisse der vorliegenden Dissertation deuten darauf hin, dass dieser Mechanismus auch bei SAD eine Rolle spielen könnte. Zusätzlich unterschieden sich die beiden Patient:innen-Gruppen mit und ohne Intrusionen in den selbst berichteten depressiven Symptomen. Mechanistisch wäre denkbar, dass depressive Symptome als Folge von belastenden Intrusionen entstehen, oder andersherum depressive Symptome das Erleben von Intrusionen begünstigen. Zusammengenommen unterstützen die Ergebnisse ein transdiagnostisches Vorgehen in Forschung und klinischer Praxis, welches sich mehr nach störungsübergreifenden Mechanismen und Symptomclustern, sowie Interventionen ausrichtet. In der klinischen Praxis bedeutet dies, dass bei Patient:innen mit SAD berücksichtigt werden sollte, ob Intrusionen in Reaktion auf ein oder mehrere ätiologisch relevante sozial aversive Ereignisse vorliegen. An die Therapie von PTBS anknüpfend könnten dann auch bei Patient:innen mit SAD diese Ereignisse in Form von sogenanntem *Reliving* (Ehlers & Clark, 2008) aufgegriffen werden (z. B. das Ereignis in der Vorstellung wiedergeben oder aufschreiben), um die Erinnerung daran zu elaborieren und damit in den richtigen Kontext einzuordnen. *Imagery Rescripting* ist ebenfalls eine Intervention, welche klassischerweise zur Behandlung traumatischer Ereignisse bei PTBS eingesetzt wird, aber auch im Kontext sozial aversiver Ereignisse bei Patient:innen mit SAD zu Erfolgen führt (Norton & Abbott, 2016a; Reimer & Moscovitch, 2015; Seinsche et al., 2023).

Zusammenfassung

Diese Dissertation untersucht die Bedrohungsverarbeitung und ihre neuronalen Korrelate bei Sozialer Angststörung (SAD), fokussiert auf zwei bisher wenig erforschte Aspekte: kontextabhängige Furchtkonditionierung und selbstbezogene Verarbeitung neutraler Gesichter. In der ersten Studie wurde die kontextabhängige Modulation von Furcht untersucht und dabei, basierend auf phänomenologischen Ähnlichkeiten zur Posttraumatischen Belastungsstörung (PTBS) erwartet, dass Patient:innen mit SAD ähnliche Defizite aufweisen wie Betroffene mit PTBS (Garfinkel et al., 2014; Milad et al., 2009; Wicking et al., 2016). Während in der Gesamtstichprobe keine eindeutigen Parallelen zur PTBS gefunden wurden,

zeigte sich in einer Untergruppe von Patient:innen mit klinisch relevanten Intrusionen in Reaktion auf ein sozial aversives Ereignis ein spezifisches Defizit: eine Hypoaktivität des ventromedialen präfrontalen Cortex (vmPFC) während des Extinktionslernens und -abrufs. Dies deutet auf eine unzureichende Nutzung von Kontextinformationen zur Furchtmodulation hin, ein bei PTBS vielfach replizierter Befund (Milad et al., 2009; Rougemont-Bücking et al., 2011). Patient:innen ohne Intrusionen wiesen hingegen andere neuronale Muster auf, was eine Differenzierung nach Subtypen nahelegt. In der zweiten Studie stand die Verarbeitung neutraler Gesichter und der Einfluss selbstbezogener Verarbeitung im Fokus. Hierzu kam ein neu entwickeltes Paradigma mit ausschließlich neutralen Gesichtern zum Einsatz. Patient:innen mit SAD bewerteten neutrale Gesichter subjektiv negativer als Kontrollpersonen, obwohl sich keine signifikanten Unterschiede in der neuronalen Aktivierung zeigten. Daraus ergibt sich, dass neutrale Gesichter keine geeignete neutrale Vergleichsbedingung, beispielsweise bei der Untersuchung der Verarbeitung emotionaler Gesichter, darstellen. Wie erwartet, führte die Verstärkung der selbstbezogenen Verarbeitung bei Patient:innen mit SAD verglichen mit gesunden Kontrollproband:innen zu einer negativeren Valenz, begleitet von einer erhöhten Aktivität in der Insula und dem ACC. Diese Befunde deuten vermutlich auf eine stärkere interozeptive Wahrnehmung (Critchley et al., 2004; Terasawa et al., 2013) und negativere selbstbezogene Gedanken (Wagner et al., 2013) in der Patient:innengruppe hin. Die Reduktion des Selbstbezugs war mit einer stärkeren Aktivierung des dlPFC verbunden, was darauf hindeutet, dass eine möglicherweise damit assoziierte Emotionsregulation in SAD mit einem höheren kognitiven Aufwand verbunden war (Koechlin et al., 2003).

Aus den Befunden lässt sich ein hypothetisches Modell ableiten, das einen wechselseitigen Zusammenhang postuliert: Defizite in der kontextabhängigen Furchtmodulation könnten zu einer erhöhten Unsicherheit in sozialen Situationen führen, was wiederum den Selbstbezug verstärkt. Dieser verstärkte Selbstbezug könnte wiederum die Nutzung externer Sicherheitssignale erschweren und so zur Aufrechterhaltung der sozialen Angst beitragen. Außerdem unterstützen die Ergebnisse der Dissertation einen transdiagnostischen Ansatz, der störungsübergreifende Mechanismen in den Fokus von

Forschung und Behandlung rückt. Therapeutisch wird empfohlen, bei Patient:innen mit SAD gezielt auslösende sozial aversive Ereignisse und damit in Zusammenhang stehende Intrusionen (z. B. in Form verzerrter Vorstellungsbilder von sich selbst) zu erfragen. Gegebenenfalls können diese Ereignisse durch Techniken wie *Reliving* oder *Imagery Rescripting* bearbeitet werden, um eine bessere Kontextualisierung zu ermöglichen. Außerdem könnte die Fähigkeit zur Unterscheidung zwischen Gefahren- und Sicherheitssignalen oder -kontexten durch Diskriminationstrainings verbessert werden.

Literatur [3]

Aderka, I. M., Hofmann, S. G., Nickerson, A., Hermesh, H., Gilboa-Schechtman, E. & Marom, S. (2012). Functional impairment in social anxiety disorder. *Journal of anxiety disorders*, *26*(3), 393–400. https://doi.org/10.1016/j.janxdis.2012.01.003

Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T. & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of anxiety disorders*, *44*, 36–46. https://doi.org/10.1016/j.janxdis.2016.10.003

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Birbaumer, N., Grodd, W., Diedrich, O., Klose, U., Erb, M., Lotze, M., Schneider, F., Weiss, U. & Flor, H. (1998). fMRI reveals amygdala activation to human faces in social phobics. *NeuroReport*, *9*(6).

Bjornsson, A. S., Hardarson, J. P., Valdimarsdottir, A. G., Gudmundsdottir, K., Tryggvadottir, A., Thorarinsdottir, K., Wessman, I., Sigurjonsdottir, Ó., Davidsdottir, S. & Thorisdottir, A. S. (2020). Social trauma and its association with posttraumatic stress disorder and social anxiety disorder. *Journal of anxiety disorders*, *72*, 102228. https://doi.org/10.1016/j.janxdis.2020.102228

Blair, K. S., Geraci, M., Devido, J., McCaffrey, D., Chen, G., Vythilingam, M., Ng, P., Hollon, N., Jones, M. & Pine, D. S. (2008). Neural response to self- and other referential praise and criticism in generalized social phobia. *Archives of general psychiatry*, *65*(10), 1176–1184. https://doi.org/10.1001/archpsyc.65.10.1176

Blair, K. S., Geraci, M., Otero, M., Majestic, C., Odenheimer, S., Jacobs, M. & Pine, D. S. (2011). Atypical modulation of medial prefrontal cortex to self-referential comments in

[3] Ausschließlich in den Artikeln zitierte Quellen befinden sich nur in den entsprechenden *References* und werden hier nicht erneut aufgenommen.

generalized social phobia. *Psychiatry research*, *193*(1), 38–45. https://doi.org/10.1016/j.pscychresns.2010.12.016

Bouton, M. E., Westbrook, R. F., Corcoran, K. A. & Maren, S. (2006). Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biological psychiatry*, *60*(4), 352–360. https://doi.org/10.1016/j.biopsych.2005.12.015

Bramson, B., Meijer, S., van Nuland, A., Toni, I. & Roelofs, K. (2023). Anxious individuals shift emotion control from lateral frontal pole to dorsolateral prefrontal cortex. *Nature communications*, *14*(1), 4880. https://doi.org/10.1038/s41467-023-40666-3

Burgio, K. L., Merluzzi, T. V. & Pryor, J. B. (1986). Effects of performance expectancy and self-focused attention on social interaction. *Journal of personality and social psychology*, *50*(6), 1216–1221. https://doi.org/10.1037//0022-3514.50.6.1216

Chauret, M., Suffren, S., Pine, D. S., Nassim, M., Saint-Amour, D. & Maheu, F. S. (2019). Fear conditioning and extinction in anxious youth, offspring at-risk for anxiety and healthy comparisons: An fMRI study. *Biological psychology*, *148*, 107744. https://doi.org/10.1016/j.biopsycho.2019.107744

Chee, M. W. L., Venkatraman, V., Westphal, C. & Siong, S. C. (2003). Comparison of block and event-related fMRI designs in evaluating the word-frequency effect. *Human brain mapping*, *18*(3), 186–193. https://doi.org/10.1002/hbm.10092

Clark, D. M. & Wells, A. (1995). A Cognitive Model of Social Phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope & F. R. Schneier (Hrsg.), *Social Phobia: Diagnosis, Assessment, and Treatment* (S. 69–93). THE GUILFORD PRESS. https://books.google.de/books?hl=de&lr=&id=rXrekuSy2bsC&oi=fnd&pg=PA69&dq=clark+wells+social+phobia+model&ots=gSdK2kuUcs&sig=fLuK2DiSmxKJIm3ZiK1A1qGDbIU#v=onepage&q=clark%20wells%20social%20phobia%20model&f=false

Cooney, R. E., Atlas, L. Y., Joormann, J., Eugène, F. & Gotlib, I. H. (2006). Amygdala activation in the processing of neutral faces in social anxiety disorder: is neutral really neutral? *Psychiatry research*, *148*(1), 55–59. https://doi.org/10.1016/j.pscychresns.2006.05.003

Cooper, S. E., van Dis, E. A. M., Hagenaars, M. A., Krypotos, A.-M., Nemeroff, C. B., Lissek, S., Engelhard, I. M. & Dunsmoor, J. E. (2022). A meta-analysis of conditioned fear generalization in anxiety-related disorders. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, *47*(9), 1652–1661. https://doi.org/10.1038/s41386-022-01332-2

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A. & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature neuroscience*, *7*(2), 189–195. https://doi.org/10.1038/nn1176

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., van den Hout, M. A. & Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and anxiety*, *32*(4), 239–253. https://doi.org/10.1002/da.22353

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

Ehlers, A. & Clark, D. M. (2008). Post-traumatic stress disorder: the development of effective psychological treatments. *Nordic journal of psychiatry*, *62 Suppl 47*(Suppl 47), 11–18. https://doi.org/10.1080/08039480802315608

Erwin, B. A., Heimberg, R. G., Marx, B. P. & Franklin, M. E. (2006). Traumatic and socially stressful life events among persons with social anxiety disorder. *Journal of anxiety disorders*, *20*(7), 896–914. https://doi.org/10.1016/j.janxdis.2005.05.006

Franke, L. K., Rattel, J. A., Miedl, S. F., Danböck, S. K., Bürkner, P.-C. & Wilhelm, F. H. (2021). Intrusive memories as conditioned responses to trauma cues: An empirically supported concept? *Behaviour Research and Therapy*, *143*, 103848. https://doi.org/10.1016/j.brat.2021.103848

Fricke, S., Seinsche, R. J., Neudert, M. K., Schäfer, A., Zehtner, R. I., Stark, R. & Hermann, A. (2024). Neural correlates of context-dependent extinction recall in social anxiety

disorder: relevance of intrusions in response to aversive social experiences. *Psychological medicine*, *54*(3), 548–557. https://doi.org/10.1017/S0033291723002179

Garfinkel, S. N., Abelson, J. L., King, A. P., Sripada, R. K., Wang, X., Gaines, L. M. & Liberzon, I. (2014). Impaired contextual modulation of memories in PTSD: an fMRI and psychophysiological study of extinction retention and fear renewal. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *34*(40), 13435–13443. https://doi.org/10.1523/JNEUROSCI.4287-13.2014

Gentili, C., Gobbini, M. I., Ricciardi, E., Vanello, N., Pietrini, P., Haxby, J. V. & Guazzelli, M. (2008). Differential modulation of neural activity throughout the distributed neural system for face perception in patients with Social Phobia and healthy subjects. *Brain research bulletin*, *77*(5), 286–292. https://doi.org/10.1016/j.brainresbull.2008.08.003

Gilboa-Schechtman, E. & Shachar-Lavie, I. (2013). More than a face: a unified theoretical perspective on nonverbal social cue processing in social anxiety. *Frontiers in human neuroscience*, *7*, 904. https://doi.org/10.3389/fnhum.2013.00904

Goldin, P. R. & Gross, J. J. (2010). Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. *Emotion (Washington, D.C.)*, *10*(1), 83–91. https://doi.org/10.1037/a0018441

Goldin, P. R., Manber, T., Hakimi, S., Canli, T. & Gross, J. J. (2009). Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. *Archives of general psychiatry*, *66*(2), 170–180. https://doi.org/10.1001/archgenpsychiatry.2008.525

Goldin, P. R., Ramel, W. & Gross, J. J. (2009). Mindfulness Meditation Training and Self-Referential Processing in Social Anxiety Disorder: Behavioral and Neural Effects. *Journal of cognitive psychotherapy*, *23*(3), 242–257. https://doi.org/10.1891/0889-8391.23.3.242

Gonzalez, S. T. & Fanselow, M. S. (2020). The role of the ventromedial prefrontal cortex and context in regulating fear learning and extinction. *Psychology & neuroscience*, *13*(3), 459–472. https://doi.org/10.1037/pne0000207

Heimberg, R. G., Brozovich, F. A. & Rapee, R. M. (2014). A Cognitive Behavioral Model of Social Anxiety Disorder. In S. G. Hofmann & P. M. DiBartolo (Hrsg.), *Social anxiety: Clinical, developmental, and social perspectives* (Third edition, S. 395–422). Academic Press is an imprint of Elsevier. https://doi.org/10.1016/B978-0-12-375096-9.00015-8

Hermann, A., Stark, R., Milad, M. R. & Merz, C. J. (2016). Renewal of conditioned fear in a novel context is associated with hippocampal activation and connectivity. *Social cognitive and affective neuroscience*, *11*(9), 1411–1421. https://doi.org/10.1093/scan/nsw047

Hermann, C., Ziegler, S., Birbaumer, N. & Flor, H. (2002). Psychophysiological and subjective indicators of aversive pavlovian conditioning in generalized social phobia. *Biological psychiatry*, *52*(4), 328–337. https://doi.org/10.1016/S0006-3223(02)01385-9

Kalisch, R., Korenfeld, E., Stephan, K. E., Weiskopf, N., Seymour, B. & Dolan, R. J. (2006). Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*(37), 9503–9511. https://doi.org/10.1523/JNEUROSCI.2021-06.2006

Kausche, F. M., Carsten, H. P., Sobania, K. M. & Riesel, A. (2025). Fear and safety learning in anxiety- and stress-related disorders: An updated meta-analysis. *Neuroscience and biobehavioral reviews*, *169*, 105983. https://doi.org/10.1016/j.neubiorev.2024.105983

Kim, S.-Y., Shin, J.-E., Lee, Y. I., Kim, H., Jo, H. J. & Choi, S.-H. (2018). Neural evidence for persistent attentional bias to threats in patients with social anxiety disorder. *Social cognitive and affective neuroscience*, *13*(12), 1327–1336. https://doi.org/10.1093/scan/nsy101

Koechlin, E., Ody, C. & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science (New York, N.Y.)*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K.,

Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of abnormal psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

Lange, W.-G., Allart, E., Keijsers, G. P. J., Rinck, M. & Becker, E. S. (2012). A neutral face is not neutral even if you have not seen it: social anxiety disorder and affective priming with facial expressions. *Cognitive behaviour therapy*, *41*(2), 108–118. https://doi.org/10.1080/16506073.2012.666563

Liberzon, I. & Abelson, J. L. (2016). Context Processing and the Neurobiology of Post-Traumatic Stress Disorder. *Neuron*, *92*(1), 14–30. https://doi.org/10.1016/j.neuron.2016.09.039

Lidle, L. R. & Schmitz, J. (2021). Rumination in Children with Social Anxiety Disorder: Effects of Cognitive Distraction and Relation to Social Stress Processing. *Research on child and adolescent psychopathology*, *49*(11), 1447–1459. https://doi.org/10.1007/s10802-021-00837-6

Lissek, S. & Grillon, C. (2012). Learning Models of PTSD. In J. G. Beck & D. M. Sloan (Hrsg.), *The Oxford Handbook of Traumatic Stress Disorders* (S. 175–190). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399066.013.0013

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S. & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *The American journal of psychiatry*, *165*(1), 124–132. https://doi.org/10.1176/appi.ajp.2007.06091513

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

Machado-de-Sousa, J. P., Arrais, K. C., Alves, N. T., Chagas, M. H. N., Meneses-Gaya, C. de, Crippa, J. A. d. S. & Hallak, J. E. C. (2010). Facial affect processing in

social anxiety: tasks and stimuli. *Journal of neuroscience methods*, *193*(1), 1–6. https://doi.org/10.1016/j.jneumeth.2010.08.013

Maren, S., Phan, K. L. & Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. *Nature reviews. Neuroscience*, *14*(6), 417–428. https://doi.org/10.1038/nrn3492

Marin, M.-F., Hammoud, M. Z., Klumpp, H., Simon, N. M. & Milad, M. R. (2020). Multimodal Categorical and Dimensional Approaches to Understanding Threat Conditioning and Its Extinction in Individuals With Anxiety Disorders. *JAMA psychiatry*, *77*(6), 618–627. https://doi.org/10.1001/jamapsychiatry.2019.4833

Marin, M.-F., Zsido, R. G., Song, H., Lasko, N. B., Killgore, W. D. S., Rauch, S. L., Simon, N. M. & Milad, M. R. (2017). Skin Conductance Responses and Neural Activations During Fear Conditioning and Extinction Recall Across Anxiety Disorders. *JAMA psychiatry*, *74*(6), 622–631. https://doi.org/10.1001/jamapsychiatry.2017.0329

Merz, C. J., Hamacher-Dang, T. C. & Wolf, O. T. (2016). Immediate extinction promotes the return of fear. *Neurobiology of learning and memory*, *131*, 109–116. https://doi.org/10.1016/j.nlm.2016.03.013

Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., Zeidan, M. A., Handwerger, K., Orr, S. P. & Rauch, S. L. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. *Biological psychiatry*, *66*(12), 1075–1082. https://doi.org/10.1016/j.biopsych.2009.06.026

Mineka, S. & Zinbarg, R. E. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *The American psychologist*, *61*(1), 10–26. https://doi.org/10.1037/0003-066X.61.1.10

Motzkin, J. C., Philippi, C. L., Wolf, R. C., Baskaya, M. K. & Koenigs, M. (2015). Ventromedial prefrontal cortex is critical for the regulation of amygdala activity in humans. *Biological psychiatry*, *77*(3), 276–284. https://doi.org/10.1016/j.biopsych.2014.02.014

Northoff, G. (2011). Self and brain: what is self-related processing? *Trends in Cognitive Sciences*, *15*(5), 186-7; author reply 187-8. https://doi.org/10.1016/j.tics.2011.03.001

Northoff, G., Heinzel, A., Greck, M. de, Bermpohl, F., Dobrowolny, H. & Panksepp, J. (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. https://doi.org/10.1016/j.neuroimage.2005.12.002

Northoff, G., Qin, P. & Feinberg, T. E. (2011). Brain imaging of the self--conceptual, anatomical and methodological issues. *Consciousness and cognition*, *20*(1), 52–63. https://doi.org/10.1016/j.concog.2010.09.011

Norton, A. R. & Abbott, M. J. (2016a). The efficacy of imagery rescripting compared to cognitive restructuring for social anxiety disorder. *Journal of anxiety disorders*, *40*, 18–28. https://doi.org/10.1016/j.janxdis.2016.03.009

Norton, A. R. & Abbott, M. J. (2016b). Self-Focused Cognition in Social Anxiety: A Review of the Theoretical and Empirical Literature. *Behaviour Change*, *33*(1), 44–64. https://doi.org/10.1017/bec.2016.2

Norton, A. R. & Abbott, M. J. (2017). Bridging the Gap between Aetiological and Maintaining Factors in Social Anxiety Disorder: The Impact of Socially Traumatic Experiences on Beliefs, Imagery and Symptomatology. *Clinical psychology & psychotherapy*, *24*(3), 747–765. https://doi.org/10.1002/cpp.2044

Ochsner, K. N. & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249. https://doi.org/10.1016/j.tics.2005.03.010

Pennington, Z. T., Anderson, A. S. & Fanselow, M. S. (2017). The ventromedial prefrontal cortex in a model of traumatic stress: fear inhibition or contextual processing? *Learning & memory (Cold Spring Harbor, N.Y.)*, *24*(9), 400–406. https://doi.org/10.1101/lm.046110.117

Peschard, V. & Philippot, P. (2017). Overestimation of threat from neutral faces and voices in social anxiety. *Journal of behavior therapy and experimental psychiatry*, *57*, 206–211. https://doi.org/10.1016/j.jbtep.2017.06.003

Phelps, E. A [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

Pineles, S. L. & Mineka, S. (2005). Attentional biases to internal and external sources of potential threat in social anxiety. *Journal of abnormal psychology*, *114*(2), 314–318. https://doi.org/10.1037/0021-843X.114.2.314

Rabinak, C. A., Mori, S., Lyons, M., Milad, M. R. & Phan, K. L. (2017). Acquisition of CS-US contingencies during Pavlovian fear conditioning and extinction in social anxiety disorder and posttraumatic stress disorder. *Journal of affective disorders*, *207*, 76–85. https://doi.org/10.1016/j.jad.2016.09.018

Rapee, R. M. & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, *35*(8), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3

Reimer, S. G. & Moscovitch, D. A. (2015). The impact of imagery rescripting on memory appraisals and core beliefs in social anxiety disorder. *Behaviour Research and Therapy*, *75*, 48–59. https://doi.org/10.1016/j.brat.2015.10.007

Rougemont-Bücking, A., Linnman, C., Zeffiro, T. A., Zeidan, M. A., Lebron-Milad, K., Rodriguez-Romaguera, J., Rauch, S. L., Pitman, R. K. & Milad, M. R. (2011). Altered processing of contextual information during fear extinction in PTSD: an fMRI study. *CNS neuroscience & therapeutics*, *17*(4), 227–236. https://doi.org/10.1111/j.1755-5949.2010.00152.x

Savage, H. S., Davey, C. G., Fullana, M. A. & Harrison, B. J. (2020). Threat and safety reversal learning in social anxiety disorder - an fMRI study. *Journal of anxiety disorders*, *76*, 102321. https://doi.org/10.1016/j.janxdis.2020.102321

Schneier, F. R., Pomplun, M., Sy, M. & Hirsch, J. (2011). Neural response to eye contact and paroxetine treatment in generalized social anxiety disorder. *Psychiatry research*, *194*(3), 271–278. https://doi.org/10.1016/j.pscychresns.2011.08.006.

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V. & Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PloS one*, *4*(6), e5865. https://doi.org/10.1371/journal.pone.0005865

Seinsche, R. J., Fricke, S., Schäfer, A., Neudert, M. K., Zehtner, R. I., Stark, R. & Hermann, A. (2023). Effects of Imagery Rescripting on Emotional Responses During Imagination of a Socially Aversive Experience. *Journal of Emotion and Psychopathology*, *1*(1), 113–128. https://doi.org/10.55913/joep.v1i1.12

Spurr, J. M. & Stopa, L. (2002). Self-focused attention in social phobia and social anxiety. *Clinical Psychology Review*, *22*(7), 947–975. https://doi.org/10.1016/S0272-7358(02)00107-1

Suarez-Jimenez, B., Albajes-Eizagirre, A., Lazarov, A., Zhu, X., Harrison, B. J., Radua, J., Neria, Y. & Fullana, M. A. (2020). Neural signatures of conditioning, extinction learning, and extinction recall in posttraumatic stress disorder: a meta-analysis of functional magnetic resonance imaging studies. *Psychological medicine*, *50*(9), 1442–1451. https://doi.org/10.1017/S0033291719001387

Terasawa, Y., Shibata, M., Moriguchi, Y. & Umeda, S. (2013). Anterior insular cortex mediates bodily sensibility and social anxiety. *Social cognitive and affective neuroscience*, *8*(3), 259–266. https://doi.org/10.1093/scan/nss108

Tinoco-González, D., Fullana, M. A., Torrents-Rodas, D., Bonillo, A., Vervliet, B., Pailhez, G., Farré, M., Andión, O., Perez, V. & Torrubia, R. (2015). Conditioned Subjective Responses to Socially Relevant Stimuli in Social Anxiety Disorder and Subclinical Social Anxiety. *Clinical psychology & psychotherapy*, *22*(3), 221–231. https://doi.org/10.1002/cpp.1883

VanElzakker, M. B., Staples-Bradley, L. K. & Shin, L. M. (2018). The Neurocircuitry of Fear and PTSD. In E. Vermetten, A. Germain & T. C. Neylan (Hrsg.), *Sleep and Combat-Related Post Traumatic Stress Disorder* (1st ed. 2018, S. 111–125). Springer New York; Imprint: Springer. https://doi.org/10.1007/978-1-4939-7148-0_10

Wagner, G., Koch, K., Schachtzabel, C., Peikert, G., Schultz, C. C., Reichenbach, J. R., Sauer, H. & Schlösser, R. G. (2013). Self-referential processing influences functional activation during cognitive control: an fMRI study. *Social cognitive and affective neuroscience*, *8*(7), 828–837. https://doi.org/10.1093/scan/nss074

Wermes, R., Lincoln, T. M. & Helbig-Lang, S. (2018). Attentional biases to threat in social anxiety disorder: time to focus our attention elsewhere? *Anxiety, stress, and coping*, *31*(5), 555–570. https://doi.org/10.1080/10615806.2018.1483497

Wicking, M., Steiger, F., Nees, F., Diener, S. J., Grimm, O., Ruttorf, M., Schad, L. R., Winkelmann, T., Wirtz, G. & Flor, H. (2016). Deficient fear extinction memory in posttraumatic stress disorder. *Neurobiology of learning and memory*, *136*, 116–126. https://doi.org/10.1016/j.nlm.2016.09.016

Wong, Q. J. J. & Rapee, R. M. (2016). The aetiology and maintenance of social anxiety disorder: A synthesis of complimentary theoretical models and formulation of a new integrated model. *Journal of affective disorders*, *203*, 84–100. https://doi.org/10.1016/j.jad.2016.05.069

Woody, S. R. & Rodriguez, B. F. (2000). Self-Focused Attention and Social Anxiety in Social Phobics and Normal Controls. *Cognitive Therapy and Research*, *24*(4), 473–488. https://doi.org/10.1023/A:1005583820758

Yoon, H.-J., Kim, J. S., Shin, Y.-B., Choi, S.-H., Lee, S.-K. & Kim, J.-J. (2016). Neural activity during self-referential working memory and the underlying role of the amygdala in social anxiety disorder. *Neuroscience letters*, *627*, 139–147. https://doi.org/10.1016/j.neulet.2016.05.068

Yoon, H.-J., Seo, E. H., Kim, J.-J. & Choo, I. H. (2019). Neural Correlates of Self-referential Processing and Their Clinical Implications in Social Anxiety Disorder. *Clinical psychopharmacology and neuroscience : the official scientific journal of the Korean College of Neuropsychopharmacology*, *17*(1), 12–24. https://doi.org/10.9758/cpn.2019.17.1.12

Yoon, K. L. & Zinbarg, R. E. (2008). Interpreting neutral faces as threatening is a default mode for socially anxious individuals. *Journal of abnormal psychology*, *117*(3), 680–685. https://doi.org/10.1037/0021-843X.117.3.680

Zhuang, Q., Yao, S., Xu, L., Chen, S., Li, J., Zheng, X., Fu, M., Kendrick, K. M. & Becker, B. (2025). A functional anatomical shift from the lateral frontal pole to dorsolateral

prefrontal cortex in emotion action control underpins elevated levels of anxiety – partial replication and generalization of Bramson et al., 2023. Vorab-Onlinepublikation. https://doi.org/10.1101/2025.03.09.642129

Danksagung

Diese Doktorarbeit wäre ohne die Unterstützung vieler Menschen nicht entstanden. Dafür möchte ich mich von ganzem Herzen bedanken.

Mein besonderer Dank gilt Apl. Prof. Dr. Andrea Hermann für die hervorragende Betreuung meiner Promotion. Ich danke ihr herzlich für ihre konstruktiven Ratschläge, ihre stete Ansprechbarkeit und dafür, dass ich sie jederzeit auf meiner Seite wähnen durfte – sowohl was das Projekt als auch alle damit verbundenen Sorgen betraf. Ihre Wertschätzung hat mich auf diesem Weg begleitet.

Herrn Prof. Dr. Rudolf Stark danke ich aufrichtig für die Möglichkeit, in seiner Abteilung zu promovieren, sowie für sein interessiertes Mitdenken und sein wertvolles Feedback.

Prof. Dr. Straube danke ich herzlich für die Übernahme der Zweitbegutachtung dieser Arbeit – insbesondere unter den gegebenen, zeitlich knappen Umständen. Allen Mitgliedern der Prüfungskommission danke ich für die aufgebrachte Zeit und Mühe.

Meinen Kolleginnen der Arbeitsgruppe, Rosa Seinsche, Raphaela Zimmer, Marie Neudert, und Ann-Kathrin Noll danke ich für die vertrauensvolle und motivierende Zusammenarbeit, das Einspringen, wenn es nötig war, und vor allem für den gemeinsamen Weg – voller Austausch, gegenseitiger Unterstützung und geteilter Momente, die auch in herausfordernden Phasen Halt und Motivation gaben.

Ein besonderer Dank gilt dem gesamten Team der Professur für Psychotherapie und Systemneurowissenschaften, bzw. des BIONs für ihre Unterstützung mit Rat und Tat. Insbesondere danke ich Dr. Axel Schäfer für seine wertvolle Unterstützung rund um die fMRT-Daten.

Allen Proband:innen, studentischen Hilfskräften und Praktikant:innen danke ich für ihre Teilnahme und Mitwirkung, die die Durchführung der Studien erst ermöglicht hat.

Zum Schluss gilt mein tiefster Dank meiner Familie. Meine Eltern haben mir stets den Rücken gestärkt und das Gefühl mit auf den Weg gegeben, dass ich meinen Weg gehen kann. Adrian danke ich für seinen Zuspruch, seinen Glauben an mich und seine Unterstützung – insbesondere im Endspurt, denn ohne ihn wäre der erfolgreiche Abschluss dieser Arbeit nicht denkbar gewesen.

Publikationen

Fricke, S., Seinsche, R. J., Neudert, M. K., Schäfer, A., Zehtner, R. I., Stark, R., & Hermann, A. (2024). Neural correlates of context-dependent extinction recall in social anxiety disorder: Relevance of intrusions in response to aversive social experiences. *Psychological Medicine*, *54*(3), 548–557. https://doi.org/10.1017/S0033291723002179

Grund, S., Breitinger, E., **Fricke, S.**, Alpers, G. W., Hundsdörfer, W., & Schafer, H. G. (2017). Health-Related Quality of Life Measurement Inpatient Geriatric Rehabilitation: A Comparison of the Feasibility and Suitability Between the SF-36 and EQ-5D-5L Questionnaires. *Journal of Gerontology & Geriatric Research*, *06*(06). https://doi.org/10.4172/2167-7182.1000458

Hermann, A., Neudert, M. K., Schäfer, A., Zehtner, R. I., **Fricke, S.**, Seinsche, R. J., & Stark, R. (2021). Lasting effects of cognitive emotion regulation: Neural correlates of reinterpretation and distancing. *Social Cognitive and Affective Neuroscience*, *16*(3), 268–279. https://doi.org/10.1093/scan/nsaa159

Neudert, M. K., Schäfer, A., Zehtner, R. I., **Fricke, S.**, Seinsche, R. J., Kruse, O., . . . Hermann, A. (2023). Behavioral pattern separation is associated with neural and electrodermal correlates of context-dependent fear conditioning. *Scientific Reports*, *13*(1), 5577. https://doi.org/10.1038/s41598-023-31504-z

Neudert, M. K., Schäfer, A., Zehtner, R. I., **Fricke, S.**, Seinsche, R. J., Stark, R., & Hermann, A. (2024). Decontextualized fear memories? Stronger conditioned fear responses during extinction learning and extinction recall in a safe context predict the development of long-term analog intrusions. *Psychological Medicine*, *54*(1), 159–168. https://doi.org/10.1017/S0033291723001125

Seinsche, R. J., **Fricke, S.**, Neudert, M. K., Zehtner, R. I., Walter, B., Stark, R., & Hermann, A. (2023). Memory representation of aversive social experiences in Social Anxiety Disorder. *Journal of Anxiety Disorders*, *94*, 102669. https://doi.org/10.1016/j.janxdis.2023.102669

Seinsche, R. J., **Fricke, S.**, Schäfer, A., Neudert, M. K., Zehtner, R. I., Stark, R., & Hermann, A. (2023). Effects of Imagery Rescripting on Emotional Responses During Imagination of a Socially Aversive Experience. *Journal of Emotion and Psychopathology*, *1*(1), 113–128. https://doi.org/10.55913/joep.v1i1.12

Seinsche, R. J., Walter, B., **Fricke, S.**, Neudert, M. K., Zehtner, R. I., Stark, R., & Hermann, A. (2022). Social phobic beliefs mediate the relationship between post-event processing regarding the worst socially aversive experience and fear of negative evaluation. *Current Psychology (New Brunswick, N.J.)*, 1–10. https://doi.org/10.1007/s12144-022-02805-9