The neurofunctional basis of affective startle modulation in humans – evidence from combined facial electromyography and functional magnetic resonance imaging

Manuel Kuhn[1*], Julia Wendt[2], Rachel Sjouwerman[1], Christian Büchel[1], Alfons Hamm[2], and Tina B. Lonsdorf[1]

[1] Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany,
[2] Department of Clinical and Physiological Psychology, University of Greifswald, Germany

Abstract

Background: The startle eye-blink is the cross-species translational tool to study defensive behavior in affective neuroscience with relevance to a broad range of neuropsychiatric conditions. It makes use of the startle reflex, a defensive response elicited by an immediate, unexpected sensory event, which is potentiated when evoked during threat and inhibited during safety. In contrast to skin conductance responses or pupil dilation, modulation of the startle reflex is valence-specific. Rodent models implicate a modulatory pathway centering on the brainstem (i.e., nucleus reticularis pontis caudalis) and the centromedial amygdala as key hubs for flexibly integrating valence information into differential startle magnitude. Technical advances now allow for the investigation of this pathway using combined facial EMG-fMRI in humans.

Methods: We employed a multi-methodological approach combining trial-by-trial facial eye-blink startle EMG and brainstem/amygdala specific fMRI in humans. Validating the robustness and reproducibility of our findings, we provide evidence from two different paradigms (fear-potentiated startle, affect-modulated startle) in two independent studies (N=43 and N=55).

Results: We provide key evidence for a conserved neural pathway for acoustic startle modulation between humans and rodents. Furthermore, we provide the crucial direct link between EMG startle eye-blink magnitude and neural response strength. Finally, we demonstrate a dissociation between arousal-specific amygdala responding and triggered valence-specific amygdala responding.

Conclusions: We provide neurobiologically-based evidence for the strong translational value of startle responding and argue that startle-evoked amygdala responding and its affective modulation may hold promise as an important novel tool for affective neuroscience and its clinical translation.

***Corresponding Author**
Manuel Kuhn
Department of Systems Neuroscience
University Medical Center Hamburg-Eppendorf
Martinistraße 52, D-20246 Hamburg, Germany
P: +49 40 7410 59364
E: m.kuhn@uke.de

Introduction

Defensive responding is innate and conserved across species with rapid protective reflexes promoting survival (1). However, ever-changing environments require flexible adaption (2). The mammalian startle reflex is elicited by an unexpected and abruptly occurring sensory stimulus (e.g. acoustic, tactile or visual) and is a prime example for the integration of short-latency responding and flexible modulation (3, 4).

In humans, the startle eye-blink reflex, as the first and most reliable component of defensive responding (5, 6), has been promoted as *the* prime cross-species translational tool for affective neuroscience with relevance to a broad range of neuropsychiatric conditions (7–16). Importantly, this responding is modulated in a valence-specific manner ('*affective startle modulation,* **ASM** (17)): decreased (*inhibited*) during positive emotional states (e.g. during viewing of positive pictures) and increased (*potentiated*) during negative emotional states (18), such as when anticipating a potential threat such as an aversive electro-tactile stimulation or during viewing of negative pictures ('*fear potentiated startle*', **FPS** (4, 7, 19)**)**. Valence-specificity of startle responding represents a major advance over other commonly employed non-valence specific measures in affective neuroscience, such as skin conductance responding (SCR) or pupil dilation (20). Yet, until recently (21, 22), technical challenges have restricted the assessment of startle responding via facial EMG recordings to behavioral studies. Hence, neurobiological models underlying this valence-dependent startle modulation are primarily derived from FPS studies in rodents (23, 24) and converge in implicating two distinct neural pathways: First, the *primary* acoustic startle pathway, conveying the startle response itself. Second, the *modulatory* pathway, adjusting response strength of the primary pathway depending on the current affective state - despite physically identical sensory input eliciting the startle response.

In rodents, the rapid *primary* acoustic startle reflex pathway involves three major hubs transferring the acoustic sensory input from the cochlear root neurons (CRNs) via the brainstem (i.e., nucleus reticularis pontis caudalis, PnC) to the motor-effectors that initiate the startle response (25, 26).

The *modulatory* pathway, which is the focus of this work, centers on the pivotal role of the PnC as the key input hub for the integration of affective modulatory information. In rodents, this modulatory input to the PnC appears to be primarily conveyed through the medial part of the central nucleus of the amygdala (24, 27–29) - the core output region initiating defensive responding (7, 30). Fine-tuning of this modulatory input is conveyed by regions exerting their influence either by modulating central amygdala activation or by direct input to the PnC (most prominently basolateral nucleus of the amygdala, BLA; bed nucleus of the stria terminalis, BNST; periaqueductal grey, PAG) (7, 24).

New technical developments now allow for combining facial EMG to assess the startle eye-blink with (f)MRI in humans and set the stage to investigate the hypothesized universality of this key defensive response pathway – the assumption underlying the promotion of startle responding as the cross-species translational tool for clinical and affective neuroscience.

Here, we comprehensively delineate the neurofunctional basis of modulatory startle responding in humans for the first time, focusing on both the PnC and the central amygdala as key structures as identified by rodent work. We assessed convergence and generality of this pathway across two well-established experimental approaches in humans: affective startle modulation (ASM) and fear-potentiated startle (FPS). Furthermore and importantly, we aim to provide a yet unexplored *direct* link between this defensive motor behavior (i.e., startle eye-blink magnitude) and neural activation to physically identical acoustic startle-probes across emotional conditions in humans.

To achieve these aims, we conducted two independent studies ($ASM_{N=43}$, $FPS_{N=55}$): First we combined the acquisition of eye-blink startle and BOLD responding as assessed via facial electromyography (EMG) and functional magnetic resonance imaging (fMRI) respectively (22, 31, 32). Second, we utilized high-resolution amygdala imaging (2mm isotropic, 6mm smoothing) as well as recent advances in human brainstem fMRI acquisition and data analysis (33, 34) supporting the investigation of the key structures expected to be involved in startle modulation (PnC, central amygdala).

Methods and Materials

Subjects and experimental design

Here, both paradigms are briefly described (for details see supplementary information (SI)).

Affective startle modulation (ASM): Forty-three male subjects [mean age (s.e.): 25.88 (0.41)] underwent a standard affective pictures startle modification task (17) including a preceding eight-trials startle habituation phase (**Figure 1B, 2A**). Twelve pictures (derived from the IAPS (35) and EmoPicS (36) databases) per emotional category (negative , neutral, positive) were selected based on matched valence and arousal ratings to elicit a reliable affective startle modulation (i.e. inhibition and potentiation, see Table S1). Post-experimental picture ratings for valence and arousal using the self-assessment manikin scale (37) were employed.

Fear-potentiated startle (FPS): Fifty-five subjects [female = 36; mean age (s.e.): 25.6 (0.47)] underwent a differential fear conditioning paradigm similar to (38) with geometric shapes as CSs (**Figure 2B**). Intermittent ratings of fear/stress/tension were acquired.

Psychophysiological data acquisition and processing

For both studies, electromyography (EMG) startle eye-blink and skin conductance responses (SCR) were acquired. Data acquisition and processing (for details see SI) were identical across studies, following published guidelines (6, 39).

Data analyses of ratings and psychophysiology

*ASM***:** Repeated-measures analyses of variance (rmANOVA) were performed in R (40)using the 'ez' package to assess differences between categories (negative, neutral, and positive) as within-subject factor for ratings of valence and arousal as well as EMG responses and SCRs

(effect sizes reported as partial η^2). Significant effects were followed up via post-hoc t-tests to specify differences across categories (for details see SI).

FPS: For ratings, EMG and SCR measures, one-sided paired-sample t-tests were performed in base R (40) for differences of mean CS+ vs. CS- responses during fear acquisition training.

Data visualizing uses the 'ggplot2' package in R.

Functional magnetic resonance Imaging (fMRI)

Data acquisition and processing

For both studies, MR data were acquired on a 3T MRI scanner (MAGNETOM Trio, Siemens, Erlangen, Germany; 12-channel head coil).

Imaging acquisition and preprocessing parameters were specifically tailored to the brainstem and amygdala for ASM, while for FPS a whole-brain approach was adopted (see SI for details).

Data analyses of the primary startle pathway

The neural response to *startle-probes* was investigated to explore the involvement of the PnC in the *primary* pathway prior to investigating the *modulatory* pathway of the startle reflex - the main focus of this work. The startle habituation phase of the ASM paradigm was particularly suited to investigate neural responding towards repetitive startle-probe presentation (**Figure 1B)** and identify PnC involvement in the primary pathway because (1) no meaningful visual stimuli are presented during this phase and (2) the timing of presentations (i.e. 11s between startle-probes plus jitter of 0, ¼, ½ or ¾ of a TR) allowed to separate the neural responses to these probes. Hence, this analysis is based on the ASM paradigm only. This analysis is based on eight habituation trials included in the first-level models described below for the valence-specific categorical analyses. Thereby, the parameter estimated for the onset reactivity for all eight probes is taken from the first-level to a one-sample t-test for second-level statistics. A directional contrast testing for positively associated activation with probe onsets was used to assess the neural responding towards the habituation startle-probes. In addition, to explore a direct brain-behavior link, EMG data were combined with the fMRI data (i.e. parametric modulation, see SI for details on methods and results).

Data analyses of the modulatory startle pathway

For both studies, a two-step approach to analyzing *neural responses to startle-probes* was employed using 1) valence-specific categorical and 2) EMG signal-integrative (i.e., parametric) analyses. First, the valence-specific categorical approach comprises *average* (i.e., across subjects) neural responses to startle-probes for all affective conditions (including subjects with insufficient EMG data quality). Second, this was complemented by analyses directly linking neural activation to the individual EMG amplitudes. Here, preprocessed trial-by-trial eye-blink data on an *individual* basis was integrated into an fMRI analyses as a parametric regressor. For all analyses we employed an *a priori* defined regions of interest approach (see SI).

Valence-specific categorical analyses: For ASM, a general linear model (GLM) included six regressors separated by stimulus category and startle condition. Stimulus presentations were

modeled as continuous blocks while overlayed startle-probes were modeled as events. Two additional regressors for the habituation startle-probes and inter-trial startle-probes were modeled as events. Additionally, one block regressor for three oddball-trials (see SI) was added. A flexible factorial design was used to, first, carry out a non-directional F-test to investigate differential neural activation to startle-probes within the regions of interest between all three valence conditions (i.e., main effect: condition). Following, *a priori* expected neural activation related to startle-probe responding during negative-valence states (i.e. startle potentiation) as compared to positive-valence states (i.e. startle inhibition) was investigated by means of a directional t-test. To explore the neural response to emotional pictures, an additional non-directional F-test based on a flexible factorial model, including estimated parameters for the emotional condition blocks, was calculated.

For FPS, a GLM included regressors for CS onsets separated by CS-type (CS+/CS-) and startle presentation (no-startle/startle) during the CS habituation as well as the fear acquisition training phase, respectively. Moreover, four additional regressors modeling the onsets of the habituation startle-probes, inter-trial startle-probes during CS habituation, inter-trial startle-probes during fear acquisition training as well as for the USs were included. Ratings across all phases were modeled in one regressor as blocks for the entire duration of each rating block. *A priori* directional t-contrasts were calculated for the hypothesized effect of interest: increased startle-probe onset reactivity during CS+ (threatening/stressful) as compared to CS- (safe/not stressful) conditions (i.e., CS+>CS-) during fear acquisition training. Second-level analysis used a one-sample t-test to test for significant differences across all individuals within the pre-defined regions of interest.

EMG signal-integrative parametric analyses: First-level models designed for integrated eye-blink response data were similar to both models used in categorical analyses for ASM and FPS (details inSI). However, for both studies, onsets for all startle-probe regressors contained in one design matrix were condensed into one single regressor of interest. To assess the correlative relationship between neural and muscular activation independent of the valence category information, recorded raw EMG magnitudes were used as parametric modulator of the startle-probe onset regressors. Second-level analyses were performed on the estimated parameters for the parametric modulator as calculated within the individual first-levels. A one-sample t-test was performed to find significant associations between neural and muscular activity.

Results

Identification of brainstem nuclei involvement in the primary acoustic startle reflex.

To investigate the neural basis of affective startle modulation, it is essential to first delineate the neural basis of the *primary* acoustic startle pathway – investigated here by utilizing the startle habituation phase which involves repetitive startle-probe presentations *without* emotional forground information in the ASM study (**Figure 1B**).

As expected from rodent work, we indeed observed activation in the PnC region ($p_{uc} < 0.001$, T = 3.47, k = 3, [x,y,z] = [2,-35,-36], **Figure 1A, Table S1**) as well as concomittant activation in secondary ROIs (i.e., CMA, PAG, both $p_{FWE(SVC)}<0.004$, **Figure 1C; Table S1)** in response

to startle-probe presentations. This supports the proposed role of the PnC as key hub in the human *primary acoustic startle reflex* pathway (see SI for additional brain-behavior correlation), which sets the stage for investigating the involvement of the proposed core regions (i.e, PnC, central amygdala) within the *modulatory startle pathway* - the main focus here.

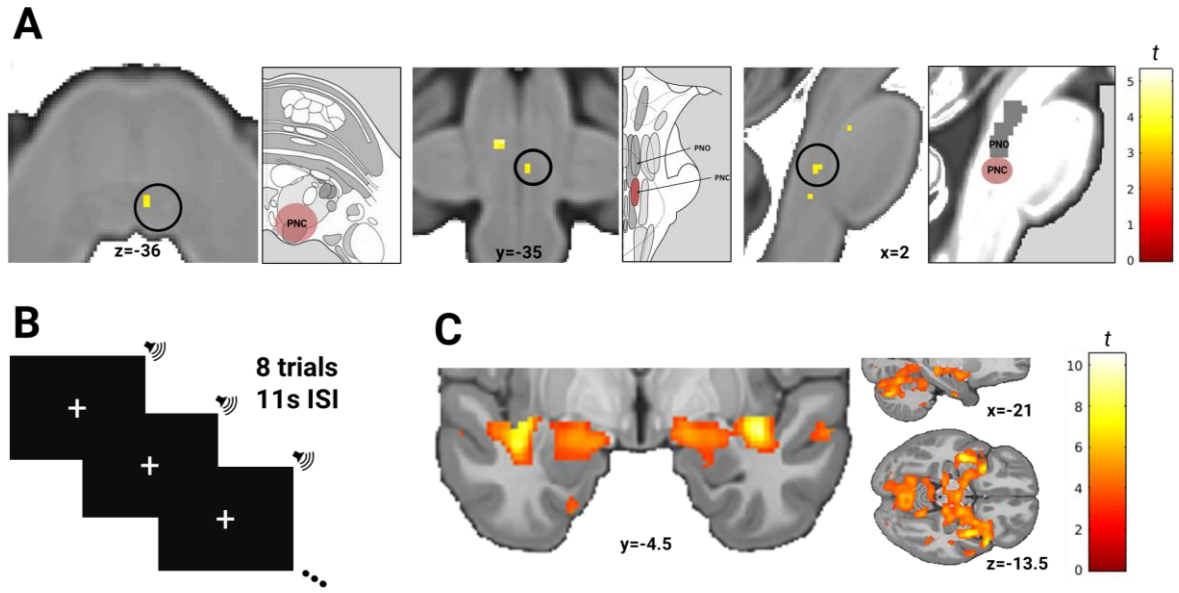

Figure 1. (A) Nucleus reticularis pontis caudalis (PnC) responding evoked by startle-probe presentations, suggesting PnC involvement in the primary acoustic startle pathway (B) During an initial startle habituation phase, eight acoustic startle-probes were presented while displaying a fixation cross as shown during the ITI. Habituation startle-probes were separated by 11s (+ added jitter). The separation of the habituation startle-probes by this long inter-stimulus-interval (ISI) and the addition of the jitter particularly allowed to quantify the individual neural response to each startle-probe (which was not possible for the short ISIs in the FPS startle-probe habituation phase. (C) Concomitant responses towards the startle-probe in the centromedial amygdala. See Figure S1A for ROI corrected visualization.

Schematic illustrations in grey Boxes in A: Location of the PnC (highlighted in red) and the nucleus pontis caudalis oralis (PnO) as defined by Duvernoy's Atlas of the Human Brain Stem and Cerebellum (left and middle, Naidich et al., 2009, adapted by permission from Springer Nature) as well as in reference to an available anatomically defined MRI ROI of the PnO (right, Edlow et al., 2012). Note that black circles highlight the activation within the PnC region and do not illustrate the specific size of search volume. Display threshold at p_{uc}<0.001. MR images are in neurological convention (left = left, right = right).

Identification of the modulatory startle pathway.

To investigate the neuro-functional basis of the *modulatory startle pathway*, we utilize two well-established experiments for affect induction to investigate a common neural pathway of affect-modulated defensive responding in humans: The *affective startle modulation* (**ASM**) paradigm and a fear conditioning paradigm allowing for the investigation of *fear-potentiated* startle (**FPS**).

On a subjective and physiological level, successful affect modulation was observed in both paradigms: In ASM, post-experimental *valence ratings* varied significantly for the three emotional picture categories [negative, neutral, positive; $F(2,84)=398.88$, $p<0.001$, $\eta^2=0.905$, **Figure 2C**] in the expected directions (one-sided: negative<neutral, negative<positive,

neutral<positive, all p<0.001). Accordingly, and replicating previous research outside the MR environment, *startle* eye-blink responses acquired during fMRI closely mirrored subjective valence ratings - commonly referred to as 'affective startle modulation' [$F(2,68)=6.29$, $p=0.003$, $\eta^2=0.156$, **Figure 2C**]. More precisely, blink magnitudes were relatively potentiated during negative (one-sided: negative>neutral: p<0.043; negative>positive: p=0.001) and inhibited during positive picture viewing (one-sided: positive<neutral: p<0.030), hence following a valence-specific gradient of startle potentiation.

In contrast, SCRs to picture onsets closely mirrored subjective *arousal ratings.* More precisely, significant differences across emotional categories [*arousal ratings*: $F(2,84)=163.74$, $p<0.001$, $\eta^2=0.796$; *SCRs:* $F(2,38)=6.31$, $p=0.004$, $\eta^2=0.223$, **Figure 2C**] reflect higher SCRs to emotionally salient (i.e., negative and positive) as compared to neutral pictures (*arousal ratings*: one-sided: negative>neutral, positive>neutral, both p<0.001; two-sided: negative vs. positive, p=0.200; *SCRs:* one-sided: negative>neutral: p<0.001, positive>neutral: p=0.003; two-sided: negative vs. neutral: p=0.541).

In FPS, successful fear acquisition was indicated by significantly higher responses to the CS+ relative to CS- across outcome measures: *fear ratings* [$t(54) = 9.55$, $p<0.001$], *startle eye-blink* [$t(50) = 2.32$, $p = 0.012$] as well as *SCRs* [$t(43) =3.62$, $p < 0.001$; **Figure 2D**].

In line with the observed valence-specific responding in subjective and psychophysiological measures, we observed stronger neural activation in PnC and CMA evoked by startle-probes presented during unpleasant (ASM: negative>positive) and threatening (FPS: CS+>CS-) conditions, which are associated with potentiated startle eye-blink responses (**Figure 3A-D, Table 1**). Of note, mirroring the valence-gradient evident from both, startle eye-blink, valence and fear ratings, PnC and CMA activation followed the same pattern (**Figure 3EF**). On a descriptive level, in both studies, amygdala activation to the startle-eliciting stimulus seems to be restricted to the dorsal part of the amygdala which, among others, includes the central nuclei (**Figure 3CD)**, – the core output area of defensive responding and proposed key effector region of the PnC.

In addition, in FPS, the BNST and PAG as our secondary ROIs were significantly implicated in fear-potentiated startle modulation (**Figure 3G, Table 1**). In ASM however, no valence-specific PAG activation was observed and the BSNT was not covered by the FOV (for details see **Table 1**).

In sum, we provide converging evidence for corresponding neural pathways underlying affect-modulated startle in rodents and humans – centering on the PnC and the CMA.

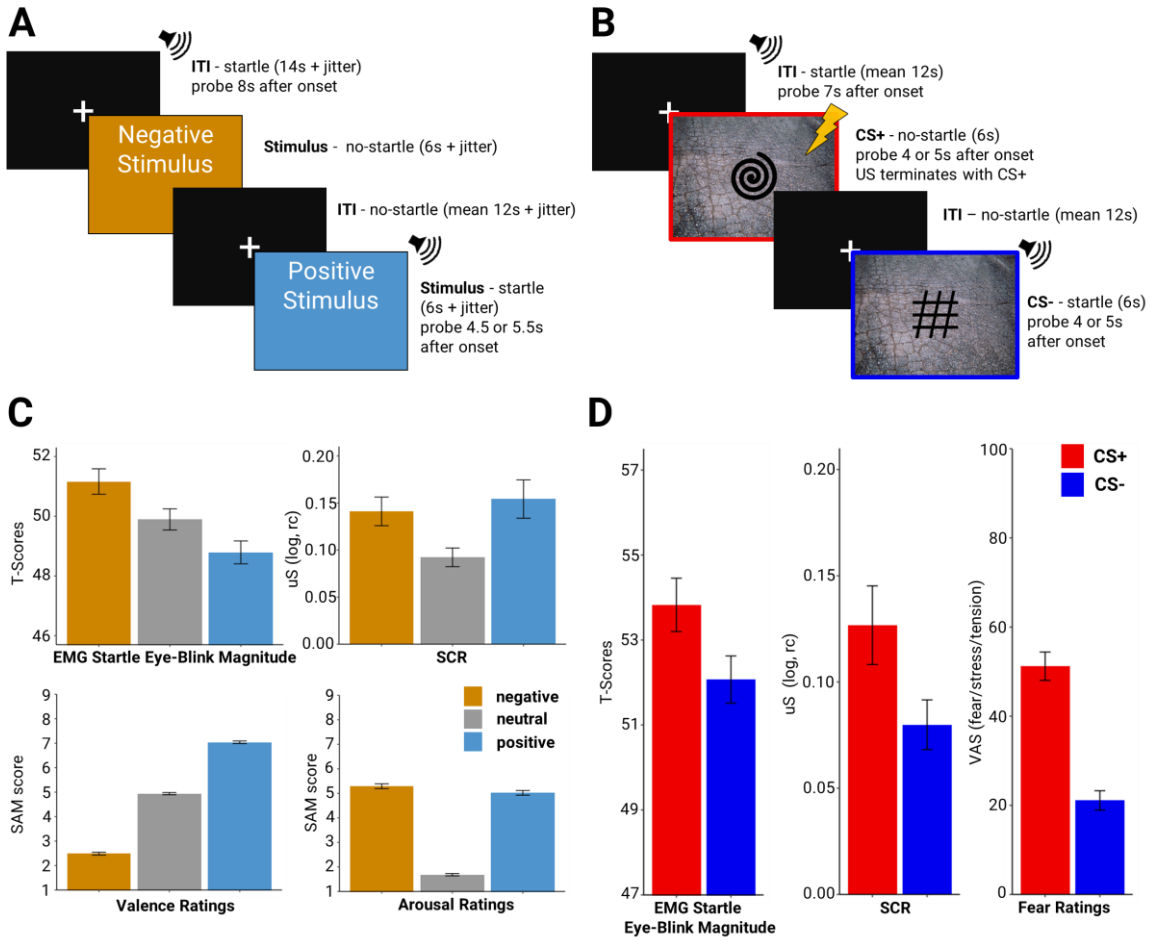

Figure 2. (A) Example of trial presentation for the affective startle modulation paradigm (ASM, neutral condition not shown). Pictures were presented for 6s separated by randomized inter-trial-intervals [ITIs, durations 10, 12 or 14s]. Moreover, a jitter (0, ¼, ½ or ¾ of a TR) added to the ITIs allowed for oversampling of the hemodynamic response function (HRF) of the BOLD-signal. Each picture was presented twice and startle-probes were presented in 50 percent of picture presentations at 4.5s or 5.5s after stimulus onset. (B) Example trial presentations for the fear conditioning (FPS) paradigm during the fear acquisition training phase (note that red and blue frames around the CS pictures serve illustrative purposes only). During fear acquisition training, CSs were presented 9 times each. One of the CSs (CS+) co-terminated with an electro-tactile US (100% reinforcement rate), whereas the other CS was never paired with an US (CS-). The startle-probe was delivered for half of the CS stimuli during CS habituation (i.e., one for CS+, one for CS), for two thirds of the CS stimuli during fear acquisition training (4 or 5s after CS onset), and for one third of all ITIs (5 or 7s after ITI onset). (C) ASM: mean responses during fMRI of startle eye-blink magnitude (affective modulation), SCR as well as post-experimental ratings. (D) FPS: mean responses during fMRI of startle eye-blink magnitude (fear-potentiated startle), SCR and subjective ratings of fear/stress/tension for CS+ and CS- in fear acquisition training. Error bars represent standard errors of the means.

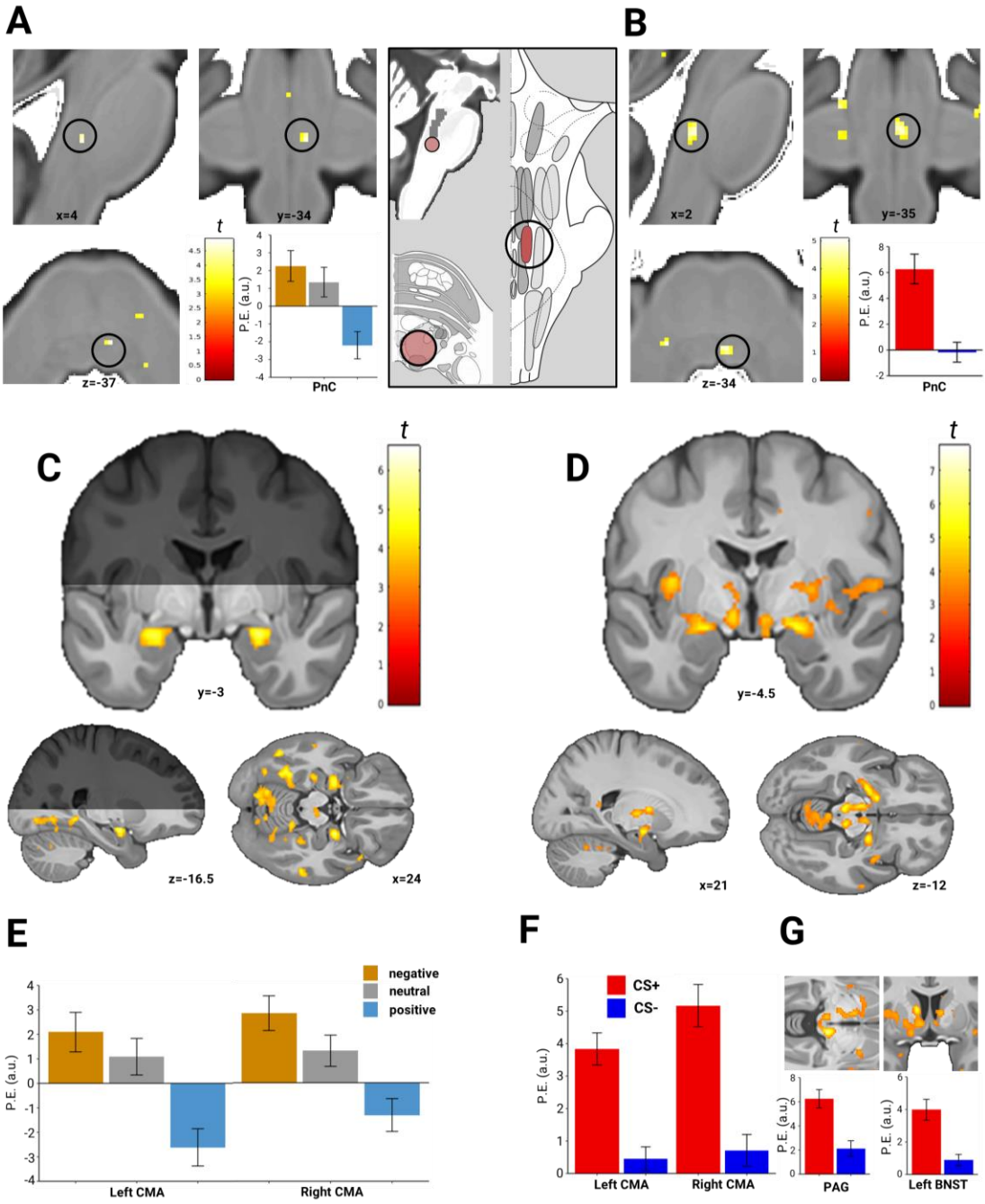

Figure 3. (A) Valence-dependent neural activation in PnC area and corresponding peak voxel parameter estimates evoked by startle-probes during ASM: negative>positive (Note that black circles highlight activation within the PnC region and does not represent a specific size of search volume). (B) Valence-dependent neural activation evoked by startle-probes in PnC area during FPS: CS+>CS- and corresponding parameter estimates of peak voxel results. See Figure S3 for an illustration of individual data points. Note that the grey box containing the schematic illustration of PnC area within Duvernoy's Atlas of the Human Brain Stem and Cerebellum (Naidich et al., 2009, modified with permission) and the available anatomically defined MRI ROI of the PnO (Edlow et al., 2012) serves to illustrate overlap between expected area of the PnC and observed statistical maps. (C) Valence-dependent neural activation in the bilateral centromedial amygdala evoked by startle-probes during ASM: negative>positive (the grey area illustrates restricted fMRI field of view) and (D) FPS: CS+>CS-. See Figure S1BC for ROI corrected visualization. (E) Corresponding parameter estimates extracted from peak voxel results in bilateral centromedial amygdala in ASM as well as (F) in FPS. (G) Valence-dependent neural activation of the PAG and the BNST in FPS and corresponding parameter estimates extracted from peak voxel. Display threshold at puc<0.001. P.E. (a.u.): parameter estimates (arbitrary units). Error bars represent standard errors of the means. MR images are in neurological convention (left = left, right = right).

Trial-by-trial brain-behavior link during affect modulation of the startle reflex.

An important further qualification of the observed valence-dependent responding on a psychophysiological and neural level can be established by *integrating* individual trial-by-trial EMG magnitudes into imaging analyses. These analyses quantify the linear relationship between EMG response magnitude and neural activation strength (i.e., parametric modulation) disregarding the categorical valence information within the statistical model.

Acquisition parameters and the design of ASM was specifically tailored to enable these methodologically challenging analyses while this question is exploratory for FPS.

In ASM, trial-by-trial magnitudes were indeed reflected in activation strength of the PnC (p_{uc} = 0.001, T = 3.65, k = 1, [x,y,z] = [2,-35,-34], **Figure 4A**) as well as left CMA ($p_{FWE(SVC)}$ = 0.014, T = 3.76, k = 5, [x,y,z] = [-24,-1.5,-16.5], **Figure 4A;** p_{SVCFWE} = 0.022, T = 3.56, k = 1, [x,y,z] = [-18,-6,-14]) providing a hitherto missing direct link between defensive behavior and corresponding neural activation.

Exploratory analyses of FPS support this association within the PnC (p_{uc} = 0.003, T = 3.20, k = 2, [x,y,z] = [5,-35,-37], **Figure 4B**) and the right CMA ($p_{FWE(SVC)}$ = 0.049, T = 3.01, k = 2, [x,y,z] = [25.5,-4.5,-12], **Figure 4B**) albeit at a more liberal threshold of p<0.005uc. Importantly, correspondence of the spatial location of significant peak voxels across both studies and analyses approaches (i.e. pre-defined categorical affective conditions vs. parametric EMG data integration) increases confidence in the observed involvement of PnC and CMA in affective startle modulation.

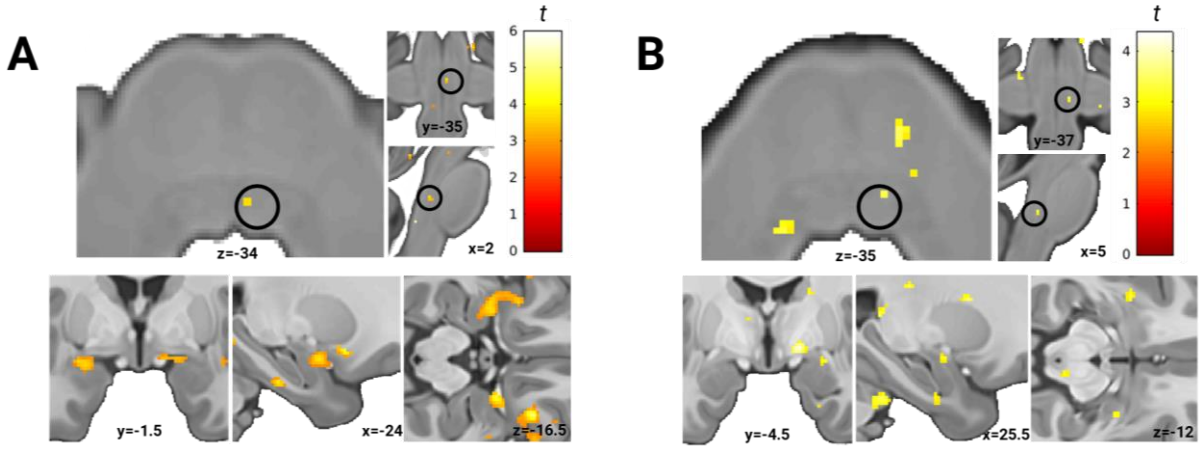

Figure 4. Activation of the PnC and CMA functionally mirroring trial-by-trial EMG magnitudes per individual in (A) ASM and (B) FPS. Display threshold at p_{uc}<0.005. See Figure S1DE for ROI corrected visualization. MR images are in neurological convention (left = left, right = right).

Dissociation in amygdala activation during passive and triggered responding

Our observation of *valence*-specific *triggered* CMA responding (i.e., evoked by the startle-eliciting stimulus, **Figure 3CE**) is intriguing, since it stands in marked contrast to the commonly observed *arousal*-dependent amygdala responding (41) during *passive* emotional picture viewing.

Importantly, investigating *passive* processing (i.e. passive viewing) of emotional pictures in our data replicates these previous reports of an arousal-dependent response pattern in SCR, arousal ratings (see **Figure 2B**), and importantly also bilateral CMA activation (**Figure 5AB**, F-test; left: $p_{FWE(SVC)} < 0.001$, F = 21.12, k = 55, [x,y,z] = [-18,-7.5,-12]; right: $p_{FWE(SVC)} < 0.001$, F = 17.07, k = 9, [x,y,z] = [19.5,-9,-13.5] and $p_{FWE(SVC)} = 0.003$, F = 10.45, k = 3, [x,y,z] = [19.5,-4.5,-15], **Figure 5**).

In sum, we observe a dissociation between centromedial amygdala responding to passive processing of emotional information (i.e. arousal-like pattern; negative and positive>neutral) and triggered centromedial amygdala responding elicited by startle-probes presented on emotional foreground information [i.e., valence specific pattern although potentiation (negative>neutral) not consistently significant].

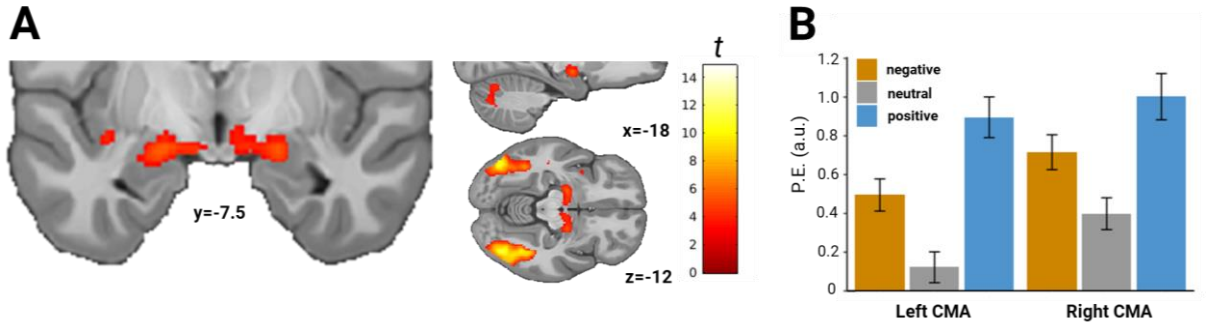

Figure 5. (A) Arousal-like pattern in the centromedial amygdala during picture viewing (emotional pictures>neutral) in the ASM study. Note that, for these analyses, only trials without startle-probes were used to avoid confounding of picture viewing-related activation by activation related to startle-probe presentation. (B) Extracted parameter estimates in left and right CMA. Display threshold at $p_{uc}<0.001$. P.E. (a.u.): parameter estimates (arbitrary units). Error bars represent standard errors of the means. MR images are in neurological convention (left = left, right = right).

Table 1. Statistics for Valence-dependent neural activation evoked by startle-probes for ASM (F-Test main effect: condition, t-contrasts for negative > positive condition) and FPS (CS+>CS-) in both a priori defined regions of interest (CMA, PnC) as well as secondary regions of interest (PAG, BNST).

Centromedial Amygdala		$p_{FWE(SVC)}$	$k_{(SVC)}$	**F**	**T**	**X**	**Y**	**Z**
Affective startle modulation (ASM)								
main effect: condition [a]	left	<0.001	60	18.55		-22	-3	-16
		0.004	5	10.43		-26	-14	-14
	right	<0.001	23	21.52		24	-3	-16
		<0.001	6	18.55		26	-4	-15
		<0.001	5	13.86		26	-12	-14
negative > positive	left	<0.001	59		4.63	-22	-4	-16
	right	<0.001	20		5.8	24	-3	-16
		0.002			4.12	20	-4	-15
		<0.001	10		4.62	26	-4	-15
Fear-potentiated startle (FPS)								
CS+ > CS-	left	<0.001	37		5.32	-24	-6	-12
	right	<0.001	18		5.37	22	-6	-12
PnC		$p_{uncorrected}$	**k**	**F**	**T**	**X**	**Y**	**Z**
Affective startle modulation (ASM)								
main effect: condition [b]		<0.001	4	13.67		4	-34	-36
		0.001	1	7.51		-3	-36	-38
negative > positive		<0.001	4		4.57	4	-34	-37
Fear-potentiated startle (FPS)								
CS+ > CS-		<0.001	24		4.96	2	-35	-34
Secondary ROIs		$p_{FWE(SVC)}$	$k_{(SVC)}$	**F**	**T**	**X**	**Y**	**Z**
PAG								
Affective startle modulation (ASM)								
main effect: condition		0.033	1	7.84		4	-30	-6
Fear-potentiated startle (FPS)								
CS+ > CS-		0.002	25		4.21	4	-30	-6
		0.012			3.64	2	-38	-10
		0.013			3.61	-3	-32	-8
		0.016			3.51	0	-34	-8
BNST								
Fear-potentiated startle (FPS)								
CS+ > CS-	left	0.001	13		4.25	-8	4	-3
	right	0.004	7		3.86	6	6	-3

[a] for completeness, we provide pair-wise comparisons based on extracted peak-voxel parameter estimates from the main effect:
Left CMA: neg vs. neu: $t(42) = 0.86$, $p = 0.397$; pos vs. neu: $t(42) = 3.56$, $p<0.001$; neg vs. pos: $t(42) = 4.53$, $p < 0.001$
Right CMA: neg vs. neu: $t(42) = 2.26$, $p = 0.029$; pos vs. neu: $t(42) = 3.12$, $p = 0.003$; neg vs. pos: $t(42) = 4.56$, $p < 0.001$

[b] for completeness, we provide pair-wise comparisons based on extracted peak-voxel parameter estimates from the main effect:
PnC: neg vs. neu: $t(42) = -0.44$, $p = 0.665$; pos vs. neu: $t(42) = 4.25$, $p<0.001$; neg vs. pos: $t(42) = 3.87$, $p < 0.001$

Discussion

The startle eye-blink reflex has been promoted as *the* prime cross-species translational tool for affective and clinical neuroscience. EMG startle responding has hitherto been employed as an additional outcome measure of *emotional processing* in the fMRI environment (21, 22, 31, 32, 42) while the neurobiological pathway underlying affective *startle responding* itself had not been investigated. Here, we utilized recent advances of combined EMG-fMRI and brainstem imaging to provide evidence for the cross-species universality of the neural pathway underlying affective startle modulation and provide the critical direct brain-behavior link across two independent samples and experimental paradigms (i.e., affective startle modulation, ASM, fear potentiated startle, FPS) in humans. In agreement with rodent work, we provide converging evidence for a conserved neural pathway centering on the PnC and the centromedial part of the amygdala (CMA). Our results further highlight the value of combining startle eye-blink EMG with fMRI measurements as a unique opportunity to probe valence-specific triggered amygdala responding as a promising novel read-out measure that can be expected to open up new avenues for affective and clinical neuroscience.

The PnC functions as key hub in the primary acoustic startle reflex (23, 24) for initiating the startle response and for integrating affective information. Here, we demonstrate startle-evoked neural responses in the PnC region also in humans. Most importantly, we show that activation in the PnC region is indeed modulated by affective input, presumably transmitted from the CMA. On a defensive response level, this manifests as affective modulation of the startle eye-blink EMG response magnitude. In addition to these key findings, we show a startle-evoked affective modulation of BNST and PAG activation in the FPS study that involved imminent threat. This corroborates their proposed involvement in the processing of fear-related information (24) and substantiates their role in defensive responding (i.e., protective reflexes such as startle), which may motivate further detailed investigations.

An important qualification of the identified affective modulation of PnC and CMA activation is the demonstration of a direct trial-by-trial brain-behavior link relating strength of neural activation to individual EMG eye-blink startle magnitudes in these key hubs of the modulating pathway. As current evidence for an association between affective (i.e., fear) modulation of the startle response has been based on lesion studies (43–45) and early PET imaging studies (46, 47), these findings provide an important direct link quantifying the relationship between eye-blink response magnitude and neural activation strength in the brainstem (i.e., PnC) as well as the CMA.

Critically, our results suggest adissociation between neural mechanisms of cue-related emotional processing and the startle reflex itself: In the behavioral lab, the dissociation between eye-blink EMG response and skin conductance responses, which mirror valence-specific and arousal-specific responding respectively (20, 48) is well described. Importantly, combining eye-blink EMG with fMRI acquisition now allowed us to demonstrate this dissociation at a neural level. In detail, we observe the expected arousal-specific CMA responding (i.e. emotional > neutral) during emotional picture viewing (41) in the ASM study, which closely follows skin conductance responses and is in line with a role of the amygdala of allocating attention to salient signals (49–52). Importantly, however, this response pattern in the CMA switches to a valence-specific responding, mirroring startle responding, through presentation of

the auditory startle-probe – an external event triggering defensive behavior. More precisely, depending on the affective state induced by the picture itself, CMA activation triggered by the startle-eliciting stimulus was either potentiated when presented on negative background information (although not consistently across brain-regions, potentially due to the male only sample in ASM(53)) or inhibited when presented on positive background information. This observation extends first hints on valence-sensitivity of the amygdala (54) and crucially supports the proposed function of the amygdala as gatekeeper for coordinated responses after initial evaluation of stimulus threat value (2). This pattern of observation is both intriguing and potentially highly relevant for future work on valence-dependent processing (55). This triggered amygdala output can be expected to mirror (observable) defensive responses towards potential threat more closely than measuring tonic amygdala responding elicited by emotional processing. Hence, such triggered events may function as a read-out of the 'state' of the amygdala, which might not be accessible otherwise. As such, we suggest that *triggered* amygdala responding may prove as a useful tool in the future.

In line with this, our results highlight the value of combining startle eye-blink EMG with fMRI measurements to provide a new (21, 22, 31, 32, 42) and, importantly, valence-specific read-out measure for affective and clinical neuroscience. Hitherto, studies have primarily used SCRs or pupil dilation in the MRI, which however, capture arousal but not valence-specific gradients (20). In particular, the observed direct relationship between startle eye-blink EMG magnitude and neural activation strength on an individual level presents a potential opportunity to use individual startle measures as direct read-out of neural activation of the central amygdala and the brainstem nuclei.

With respect to its clinical application, our work may set the ground for in-depth examination of the neural mechanisms underlying previous reports relating the modulation of the startle reflex to effects of psychopharmacology, genetics and, importantly, psychiatric conditions (7–12, 56–58).

Some limitations of our work are, however, worth noting: First, defining the exact anatomical location of most brainstem nuclei is a challenge as anatomically defined boundaries are not available. Consequently, our brainstem fMRI results are based on uncorrected thresholds and hence must be considered rather preliminary. Yet, we provide both spatially and functionally converging evidence from two independent samples and paradigms that support the accuracy of the PnC location in our work. Hence, we provide anatomical coordinates that future work may utilize for defining the PnC area.

Second, it can hence only be speculated that our results using acoustic triggers generalize to other trigger modalities (e.g., tactile, visual).

Third, our work primarily focused on potentiation of the startle reflex while inhibition of the startle reflex might also be of interest for future work.

Forth, we demonstrate a general involvement of the PnC and CMA in the affective modulation of the startle reflex. Future studies targeting the specific interconnections between both areas (e.g. by using functional connectivity analyses) are warranted to explore the mechanisms underlying startle modulation in more detail.

Fifth, analyses of the primary startle reflex are based on a limited number of trials (eight) which might lead to unreliable parameter estimates.

In conclusion, in human affective neuroscience, reflexive responding and its adaptation to environmental demands has hitherto not received much attention (however, see (59)) - in contrast to higher order (cognitive) components of emotional processing and regulation. By highlighting the cross-species conserved neural pathway of defensive startle reflex modulation, we provide an important yet missing piece connecting hitherto separate lines of research on 1) the role of the amygdala in emotion processing in humans (e.g. fear learning) and 2) the role of the amygdala in affective startle reflex modulation in rodents. This corroborates the role of startle reflex modulation as *the* prime cross-species translational tool of defensive reactivity in clinical and affective neuroscience (15, 60). This is reflected in startle potentiation being incorporated in the RDoC matrix under the acute ("fear") and potential ("anxiety") threat construct (61, 62). Critically however, its application in humans has been limited to behavioral work by technical and methodological constraints in the past. Here, we demonstrate both the applicability of EMG eye-blink startle responding in the fMRI context and provide the crucial direct brain-behavior link for affective startle modulation. This will allow to explore entirely new avenues in the future that can be expected to provide major novel insights in affective neuroscience.

Acknowledgements/ Disclosures

The authors thank Christian Möller, Jürgen Finsterbusch and Katja Hillbrandt for technical support for preparing EMG-fMRI data acquisition, Maike Möller and Jana Hofacker for help with participant preparation and data acquisition, Christian Sprenger for help with data analysis, Katrin Bergholz and Kathrin Wendt for technical assistance during MR data acquisition, as well as Jan Haaker and Christoph Korn for comments on previous versions of this manuscript.

The author thank the University Medical Center Hamburg Eppendorf (Forschungsfond Medizin), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projectnumber 44541416 – TRR 58 (sub-project B07) to TBL.

Data availability

FMRI group statistics (T- and F-maps) of all analyses from ASM and FPS presented within the main text are available on Neurovault for download: https://neurovault.org/collections/4469/.

Behavioral and psychophysiological data is available upon request.

References

1. LeDoux J (2012): Rethinking the emotional brain. *Neuron*. 73: 653–76.

2. Mobbs D, Hagan CC, Dalgleish T, Silston B, Prévost C (2015): The ecology of human fear: Survival optimization and the nervous system. *Front Neurosci*. 9.

3. Brown P, Rothwell JC, Thompson PD, Britton TC, Day BL, Marsden CD (1991): New observations on the normal auditory startle reflex in man. *Brain*. 114: 1891–1902.

4. Hamm AO (2015): Fear-Potentiated Startle. *Int Encycl Soc Behav Sci*. pp 860–867.

5. Anthony BJ (1985): In the blink of an eye: Implications of reflex modification for information processing. In: Ackles PK, Jennings JR, Coles MGH, editors. *Adv Psychophysiol*. Greenwich, CT, pp 167–218.

6. Blumenthal TD, Cuthbert BN, Filion DL, Hackley S, Lipp O V, van Boxtel A (2005): Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*. 42: 1–15.

7. Davis M, Walker DL, Miles L, Grillon C (2010): Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology*. 35: 105–35.

8. Romero M, Williams WC, New AS, Siever LJ, Speiser LJ, Hazlett EA, *et al.* (2007): Exaggerated Affect-Modulated Startle During Unpleasant Stimuli in Borderline Personality Disorder. *Biol Psychiatry*. 62: 250–255.

9. Allen NB, Trinder J, Brennan C (1999): Affective startle modulation in clinical depression: Preliminary findings. *Biol Psychiatry*. 46: 542–550.

10. Quednow BB, Frommann I, Berning J, Kühn KU, Maier W, Wagner M (2008): Impaired Sensorimotor Gating of the Acoustic Startle Response in the Prodrome of Schizophrenia. *Biol Psychiatry*. 64: 766–773.

11. Conzelmann A, Mucha RF, Jacob CP, Weyers P, Romanos J, Gerdes ABM, *et al.* (2009): Abnormal Affective Responsiveness in Attention-Deficit/Hyperactivity Disorder: Subtype Differences. *Biol Psychiatry*. 65: 578–585.

12. Schmidt U, Kaltwasser SF, Wotjak CT (2013): Biomarkers in posttraumatic stress disorder: Overview and implications for future research. *Dis Markers*. 35: 43–54.

13. Moberg CA, Bradford DE, Kaye JT, Curtin JJ (2017): Increased startle potentiation to unpredictable stressors in alcohol dependence: Possible stress neuroadaptation in humans. *J Abnorm Psychol*. 126: 441–453.

14. Winslow JT, Parr LA, Davis M (2002): Acoustic startle, prepulse inhibition, and fear-potentiated startle measured in rhesus monkeys. *Biol Psychiatry*. 51: 859–866.

15. Glover EM, Phifer JE, Crain DF, Norrholm SD, Davis M, Bradley B, *et al.* (2011): Tools for translational neuroscience: PTSD is associated with heightened fear responses using acoustic startle but not skin conductance measures. *Depress Anxiety*. 28: 1058–1066.

16. Jovanovic T, Blanding NQ, Norrholm SD, Duncan E, Bradley B, Ressler KJ (2009): Childhood abuse is associated with increased startle reactivity in adulthood. *Depress Anxiety*. 26: 1018–1026.

17. Lang PJ, Bradley MM, Cuthbert BN (1990): Emotion , Attention , and the Startle Reflex. *Psychol Rev*. 97: 377–395.

18. Davis M, Walker DL, Lee Y (1999): Neurophysiology and neuropharmacology of startle and its affective modulation. In: Dawson ME, Schell AM, Bohmelt AH, editors. *Startle Modif - Implic Neurosci Cogn Sci Clin Sci*. Cambridge University Press, pp 95–113.

19. Hamm AO, Vaitl D (1996): Affective learning: Awareness and aversion. *Psychophysiology*. 33: 698–710.

20. Bradley MM, Miccoli L, Escrig MA, Lang PJ (2008): The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*. 45: 602–607.

21. van Well S, Visser RM, Scholte HS, Kindt M (2012): Neural substrates of individual differences in human fear learning: evidence from concurrent fMRI, fear-potentiated startle, and US-expectancy data. *Cogn Affect Behav Neurosci*. 12: 499–512.

22. Lindner K, Neubert J, Pfannmöller J, Lotze M, Hamm AO, Wendt J (2015): Fear-potentiated startle processing in humans: Parallel fMRI and orbicularis EMG assessment during cue conditioning and extinction. *Int J Psychophysiol*. 98: 535–545.

23. Yeomans JS, Frankland PW (1995): The acoustic startle reflex: neurons and connections. *Brain Res Rev*. 21: 301–314.

24. Koch M (1999): The neurobiology of startle. *Prog Neurobiol*. 59: 107–128.

25. Lee Y, López DE, Meloni EG, Davis M (1996): A primary acoustic startle pathway: obligatory role of cochlear root neurons and the nucleus reticularis pontis caudalis. *J Neurosci*. 16: 3775–3789.

26. Gómez-Nieto R, de Horta-Júnior J de AC, Castellano O, Millian-Morell L, Rubio ME, López DE (2014): Origin and function of short-latency inputs to the neural substrates underlying the acoustic startle reflex. *Front Neurosci*. 8: 216.

27. Hitchcock JM, Davis M (1986): Lesions of the amygdala, but not of the cerebellum or red nucleus, block conditioned fear as measured with the potentiated startle paradigm. *Behav Neurosci*. 100: 11–22.

28. Hitchcock JM, Davis M (1991): Efferent pathway of the amygdala involved in conditioned fear as measured with the fear-potentiated startle paradigm. *Behav Neurosci*. 105: 826–842.

29. Rosen JB, Hitchcock JM, Sananes CB, Miserendino MJD, Davis M (1991): A Direct Projection From the Central Nucleus of the Amygdala to the Acoustic Startle Pathway: Anterograde and Retrograde Tracing Studies. *Behav Neurosci*. 105: 817–25.

30. LeDoux JE, Iwata J, Cicchetti P, Reis DJ (1988): Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *J*

Neurosci. 8: 2517–2529.

31. de Haan MIC, van Well S, Visser RM, Scholte HS, van Wingen GA, Kindt M (2018): The influence of acoustic startle probes on fear learning in humans. *Sci Rep*. 8: 14552.

32. Wendt J, Löw A, Weymar M, Lotze M, Hamm AO (2017): Active avoidance and attentive freezing in the face of approaching threat. *Neuroimage*. 158: 196–204.

33. Sclocco R, Beissner F, Bianciardi M, Polimeni JR, Napadow V (2018): Challenges and opportunities for brainstem neuroimaging with ultrahigh field MRI. *Neuroimage*. 168: 412–426.

34. Beissner F (2015): Functional MRI of the Brainstem: Common Problems and their Solutions. *Clin Neuroradiol*. 25: 251–257.

35. Lang PJ, Bradley MM, Cuthbert BN (1997): International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *NIMH Cent Study Emot Atten*. .

36. Wessa M, Kanske P, Neumeister P, Bode K, Heissler J, Schönfelder S (2010): EmoPics: Subjektive und psychophysiologische Evaluationen neuen Bildmaterials für die klinisch-bio-psychologische Forschung. *Zeitschrift für Klin Psychol und Psychother*. 1/11.

37. Bradley MM, Lang PJ (1994): Measuring emotion: The self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 25: 49–59.

38. Sjouwerman R, Niehaus J, Kuhn M, Lonsdorf TB (2016): Don't startle me—Interference of startle probe presentations and intermittent ratings with fear acquisition. *Psychophysiology*. 53: 1889–1899.

39. Boucsein W, Fowles DC, Grimnes S, Ben-Shakhar G, Roth WT, Dawson ME, Filion DL (2012): Publication recommendations for electrodermal measurements. *Psychophysiology*. 49: 1017–1034.

40. R Developement Core Team (2015): R: A Language and Environment for Statistical Computing. *R Found Stat Comput*. 1: 409.

41. Costa VD, Lang PJ, Sabatinelli D, Bradley MM, Keil A (2009): The Timing of Emotional Discrimination in Human Amygdala and Ventral Visual Cortex. *J Neurosci*. 29: 14864–14868.

42. Heller AS, Lapate RC, Mayer KE, Davidson RJ (2014): The face of negative affect: trial-by-trial corrugator responses to negative pictures are positively associated with amygdala and negatively associated with ventromedial prefrontal cortex activity. *J Cogn Neurosci*. 26: 2102–10.

43. Weike AI, Hamm AO, Schupp HT, Runge U, Schroeder HSW, Kessler C (2005): Fear Conditioning following Unilateral Temporal Lobectomy: Dissociation of Conditioned Startle Potentiation and Autonomic Learning. *J Neurosci*. 25: 11117–11124.

44. Klumpers F, Morgan B, Terburg D, Stein DJ, van Honk J (2014): Impaired acquisition of classically conditioned fear-potentiated startle reflexes in humans with focal bilateral basolateral amygdala damage. *Soc Cogn Affect Neurosci*. 10: 1161–1168.

45. Angrilli A, Mauri A, Palomba D, Flor H, Birbaumer N, Sartori G, Di Paola F (1996): Startle reflex and emotion modulation impairment after a right amygdala lesion. *Brain*. 119: 1991–2000.

46. Pissiota A, Frans O, Fredrikson M, Langstrom B, Flaten MA (2002): The human startle reflex and pons activation: a regional cerebral blood flow study. *Eur J Neurosci*. 15: 395–398.

47. Pissiota A, Frans Ö, Michelgård Å, Appel L, Långström B, Flaten MA, Fredrikson M (2003): Amygdala and anterior cingulate cortex activation during affective startle modulation: A PET study of fear. *Eur J Neurosci*. 18: 1325–1331.

48. Balaban MT, Taussig HN (1994): Salience of fear/threat in the affective modulation of the human startle blink. *Biol Psychol*. 38: 117–131.

49. Dal Monte O, Costa VD, Noble PL, Murray EA, Averbeck BB (2015): Amygdala lesions in rhesus macaques decrease attention to threat. *Nat Commun*. 6.

50. Dolan RJ, Vuilleumier P (2006): Amygdala Automaticity in Emotional Processing. *Ann N Y Acad Sci*. 985: 348–355.

51. Vuilleumier P, Pourtois G (2007): Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*. 45: 174–194.

52. Wendt J, Weike AI, Lotze M, Hamm AO (2011): The functional connectivity between amygdala and extrastriate visual cortex activity during emotional picture processing depends on stimulus novelty. *Biol Psychol*. 86: 203–209.

53. Stevens JS, Hamann S (2012): Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia*. 50: 1578–1593.

54. Anders S, Eippert F, Weiskopf N, Veit R (2008): The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: An fMRI study. *Soc Cogn Affect Neurosci*. 3: 233–243.

55. Tye KM (2018): Neural Circuit Motifs in Valence Processing. *Neuron*. 100: 436–452.

56. Calder AJ, Goodyer IM, Stobbe Y, van Goozen SHM, Fairchild G (2010): Facial Expression Recognition, Fear Conditioning, and Startle Modulation in Female Subjects with Conduct Disorder. *Biol Psychiatry*. 68: 272–279.

57. Patrick CJ, Berthot BD, Moore JD (1996): Diazepam blocks fear-potentiated startle in humans. *J Abnorm Psychol*. 105: 89–96.

58. Deckert J, Weber H, Villmann C, Lonsdorf TB, Richter J, Andreatta M, *et al.* (2017): GLRB allelic variation associated with agoraphobic cognitions, increased startle response and fear network activation: a potential neurogenetic pathway to panic disorder. *Mol Psychiatry*. 22: 1431–1439.

59. Roelofs K (2017): Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philos Trans R Soc B Biol Sci*. 372.

60. Hamm AO, Richter J, Pané-Farré C, Westphal D, Wittchen H-UU, Vossbeck-Elsebusch AN, *et al.* (2016): Panic disorder with agoraphobia from a behavioral neuroscience perspective: Applying the research principles formulated by the Research Domain Criteria (RDoC) initiative. *Psychophysiology*. 53: 312–322.

61. Insel TR, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, *et al.* (2010): Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 167: 748–751.

62. Lonsdorf TB, Richter J (2017): Challenges of fear conditioning research in the age of RDoC. *Zeitschrift fur Psychol / J Psychol*. 225: 189–199.

63. Naidich TP, Duvernoy HM, Delman BN, Sorensen AG, Kollias SS, Haacke EM (2009): *Duvernoy's Atlas of the Human Brain Stem and Cerebellum: High-Field MRI, Surface Anatomy, Internal Structure, Vascularization and 3 D Sectional Anatomy*. Springer Science & Business Media.

64. Edlow BL, Takahashi E, Wu O, Benner T, Dai G, Bu L, *et al.* (2012): Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. *J Neuropathol Exp Neurol*. 71: 531–46.

Supplementary Information

Supplementary Methods and Materials

Subjects

Subjects were recruited via online advertisement and provided written informed consent. Protocols were approved by the ethics commission of the German Psychological Society (DGPs) (**affective startle modulation**, **ASM**) or the General Medical Council Hamburg (**fear-potentiated startle**, **FPS**).

Experimental design

Affective startle modulation (ASM): Only male subjects were included to avoid sex-specific stimulus selections since erotic pictures served as the positive stimulus category. See **Table S1** for stimuli numbers and mean valence and arousal values as indicated by the databases manual. The paradigm was validated in a preceding behavioral pilot study in an independent sample (N_{pilot} = 24, data not shown). Pictures with comparable social content were selected. Time-points of startle probe presentations were selected to optimize expected differences in startle response magnitudes between valence categories, which were shown to be maximal in late phases of picture processing (1, 2), as well as allowing for a combined assessment of startle responses and skin conductance responses (SCRs) to visual stimuli that are unaffected by SCRs elicited by the startle probe itself. Based on this, startle probes were presented either at 4.5s or 5.5s in a counterbalanced fashion across categories and trial lists. To avoid predictability of the startle probes, 12 startle probes were added across ITIs each occurring during one of the 14s ITI (+added jitter) periods with onset 8s after ITI onset.

To increase subjects' alertness, an 'oddball task' was included. Subjects were instructed to press a button whenever a scrambled picture was presented. These pictures were taken from the neutral picture group, scrambled in cubes (25-by-25 pixels in size) and not recognizable in content.

In total, each picture per category was presented twice with three additional oddball presentations resulting in 75 trials.

Pictures (800x600 pixels) presented on a grey background were projected onto a screen (1024x786 pixels) at the back of the magnet's bore within the MR scanner which participants could see via a mirror mounted over their heads. Visual and auditory stimuli were presented using Psychophysics Toolbox-3 (5) running on MATLAB2010b (The MathWorks, Natick, MA, USA).

Fear-potentiated startle (FPS): The fear conditioning paradigm [experimentally similar to Sjouwerman et al. (2016)(6), **Figure 2B**] consisted of six phases: startle probe habituation to achieve a stable baseline for startle reactivity, CS habituation, fear acquisition training, immediate extinction, reinstatement, and a reinstatement test. The present study focuses on the startle probe habituation, CS habituation and fear acquisition training phases only and hence we provide no further details with regard to the other experimental phases. Two geometric shapes (hash and spiral, **Figure 2B**) on a experimentally unrelated background (water, sand,

grass or concrete) served as CSs. Allocation of the CS to the CS+ and CS- and the order in which the CS+/CS- appeared were counterbalanced between individuals. Prior to the experiment, participants were explicitly instructed to not attend to the startle probes to avoid interference with CS-US contingency acquisition. No explicit instructions regarding the CS-US contingencies were provided.

Table S1. Selected stimuli used for the affective startle modulation paradigm (ASM) taken from the IAPS (3) and EmoPics (4) databases. Provided are stimuli numbers for each emotional picture database and details on mean valance and mean arousal ratings as indicated by the databases manuals per emotional condition.

Selected Stimuli	**Negative**	**Neutral**	**Positive**
IAPS #'s	3225, 6312, 6315, 6560, 6250, 6313, 6510, 6230	2036, 2393,2396, 2026	4641, 4680, 4658, 4695, 4697, 4002, 4659, 4085
EmoPics #'s	215, 249, 248, 242	171, 125, 166, 160, 119, 111, 124, 127	050, 051, 052, 054
Mean Valence	2,53	5,00	7,43
Mean Arousal	6.50	2.99	6.49

During startle probe habituation, five startle probes (ISI: 6s) were presented while displaying a white fixation cross on black background which was also shown during the inter-trial interval (ITI, durations 10, 11, 12 or 13s) during the following experimental phases. During CS stimuli habituation, both CSs were presented twice. The first and last CS presentations per trial type during fear acquisition training were always presented with a startle probe. As for the ASM study, time points of startle probe onsets (4s or 5s post CS-onset) were selected to optimize differences in startle responses across CSs (1, 2) and combined measurements with SCR as well as to avoid interference with unconditioned SCR responding (i.e., to the US and startle probe). Presentation of all stimuli was controlled using Presentation Software (NeuroBehavioral Systems, Albany, CA). Visual stimuli were projected onto a screen (1024x786 pixels) at the back of the magnet's bore within the MR scanner which participants could see via a mirror mounted over their heads.

The unconditioned stimulus (US) was administered as an electro-tactile stimulus consisting of a train of three 2-ms square waves with an ISI of 50ms to the back of the right hand. The electrical stimulation was generated by a DS7A electrical stimulator (Digitimer, Welwyn Garden City, UK, delivering outputs up to 100mA) and delivered through an electrode with a platinum pin surface (Specialty Developments, Bexley, UK). Prior to the experiment, US intensity was calibrated for each participant individually [unpleasant but tolerable, aiming at an intensity of 7 out of 10 (with 10 referring to the most unpleasant sensation that might be

inducted by the electrode), mean intensity (s.e.) of the final sample (N=55): 5.12mA (0.47)]. Intermittent ratings of fear/stress/tension were acquired (see below for details).

Subjective ratings

*ASM***:** Outside the MR-environment subjects rated each picture at a computer screen (screen size 1920x1200 pixels, stimulus size 800x600 pixels) after scanning. A 9-Point Self-Assessment-Manikin [SAM (7) rating scale for valence (from 1 = very pleasant to 9 = very unpleasant) and arousal (from 1 = very calm to 9 = very arousing)] was used. After one initial training trial using a novel neutral stimulus to familiarize the subject with the rating procedure, all pictures from all three categories were consecutively presented at random with rating scales of valence and arousal, respectively, beneath the picture. Ratings were selected via mouse click at the subject's own pace and rating times were recorded to check for compliance of the subject.

*FPS***:** Participants indicated their level of fear, anxiety, and distress toward both CS types ("How much stress, fear, or anxiety did you experience the last time you saw symbol X?" with the X referring to one of the CS types at a time) intermittently throughout the experiment (in one rating block after CS habituation and three rating blocks during fear acquisition training) on a visual analogue scale (VAS) ranging from 0 (none) to 100 (maximum). A rating block was preceded by a screen that signaled the start of the rating block for 4s. The rating block included one rating for the CS+ and one for the CS- in a randomized order. The start position of the curser was randomly placed on the VAS for every trial. Participants were required to confirm their rating within 9s otherwise the rating trial was regarded as invalid and treated as missing. Ratings within blocks were separated by an ITI of 1s. An additional rating for the aversiveness of the startle sound was included after the habituation phase.

Psychophysiological data acquisition and processing

For both studies, electromyography (EMG) startle eye-blink and skin conductance responses (SCR) were acquired. Data acquisition and processing were identical across studies.

EMG: In both studies, startle eye-blink response data were acquired through EMG recordings within the MR environment using a FaceEMG Cap-MR (EasyCap GmbH, Herrsching, Germany). The cap contains five Ag/AgCl electrodes and built-in 5 kOhm resistors to facilitate wire management in the magnetic field, to ensure participant and equipment safety as well as artifact reduction in EMG and BOLD data. Skin was prepared with abrasive electrode gel. Two electrodes were placed at the participant's orbicularis oculii muscle beneath the left eye; two electrocardiogram (ECG) electrodes were placed at the participant's back and one ground electrode was attached to the forehead of the participant. The maximum transmission resistance threshold was set to 20 kOhm.

A 50ms burst of white noise served as startle probe which was calibrated to 103dB[A] using a MR-compatible sound level meter (Optoacoustics Ltd., Mazor, Israel) and presented to subjects binaurally via headphones (MR Confon GmbH, Magdeburg, Germany). Baseline scanner noise during EPI acquisition in both studies was approximately 79dB[A]. EMG data were recorded within the BrainVision Recorder software using BrainAmp ExG MR amplifier (Brain Products GmbH, Gilching, Germany), including the SyncBox device for synchronization of recorded

data and MR gradient switching (8), applying a 16 bit Analog-to-Digital-Conversion (ADC). Sampling rate was set to 5 kHz with a signal resolution of 0.1 μV within a frequency band of 0.016 and 250 Hz.

Data were processed as described previously (9). Briefly, this included MR gradient correction by subtraction of an artifact template averaged over seven EPI volumes from the raw signal, down-sampling to 1000 Hz, reduction of eye-movement related artifacts by application of a low cutoff filter of 60 Hz with a time constant of 0.0027 and 48 db/oct, rectification, and manual offline scoring using a custom-made computer program. Following published guidelines (10), the magnitude of the eye-blink response (in microvolts) was measured from onset to peak, as described previously (11). Eye-blink magnitudes were T-transformed (including all experimental phases and conditions, see below) for statistical analyses of startle responses while raw values were fed into fMRI trial-by-trial first-level analyses (see fMRI analyses for details). Undetectable blinks were scored as zero responses and as missing if a blink occurred immediately (up to 50ms) before startle probe administration or due to excessive baseline activity, obvious electrode, or gradient artefacts.

SCR: In both studies, SCRs were measured via self-adhesive Ag/AgCl electrodes which were placed on the palmar side of the left hand on the distal and proximal hypothenar. Hands were washed with tap water and without soap. Data were recorded with a CED2502-SA skin conductance unit together with a Biopac MP150-amplifier system (BIOPAC Systems Inc, Goleta, California, USA) with Spike 2 software (Cambridge Electronic Design, Cambridge, UK). Data were down-sampled to 10 Hz, smoothed by using a 5-point moving average and phasic SCR to stimulus onsets were manually scored offline using a custom-made computer program. SCR amplitudes (in μS) were scored as the largest response initiating 0.9 to 4.0 s after stimulus onset (12). Non-responses were scored as zero and trials showing recording artefacts were scored as missing data. Logarithms were computed for all values to normalize the distribution (13), and these log values were range-corrected (SCR/SCRmax) to account for inter-individual variability (14).

Subject preparation and EMG data quality control

In both studies in the scanner, SCR electrodes and a respiration belt were attached, a pulse-oximeter was attached to the left index finger and headphones were placed on the subject's head. Afterwards, the EMG cap was connected to the EMG amplifier. We ensured no heat build-up in the electrodes and that the subject's visual area was not restricted by the EMG equipment placed behind the projection screen. Subsequently, impedances of the EMG/ECG-electrodes were re-checked to ensure subject's safety. SCR, pulse and respiration signals were visually inspected to check for data quality of physiological responses prior to the experiment. The auditory startle probe was presented to ensure the subject's compliance with the sound level of the stimulus and to check data quality of the EMG signal without gradient artifacts of the scanner (i.e. scanner offline). Following this set-up procedure, a structural image (see MRI acquisition) was acquired to allow the subject to get used to the environment and to ensure all equipment was working safely. The subject was reminded that from time to time an auditory stimulus will be presented without any relevance to the stimuli. The subjects was informed that the experiment will begin with the presentation of several (ASM: eight / FPS: five) of the

auditory stimuli before the first visual stimuli are presented (i.e. startle probe habituation phase). Starting with the scanner gradient, the EMG signal was visually inspected via online MR-artifact correction with Brain Vision's RecView to ensure all electrodes were still attached and signal quality was good.

Data analyses of ratings and psychophysiology

Data analyses were homogenized across both studies whenever feasible. Insufficient data quality (defined by more than 66% of missing values or null responses for ASM: during startle habituation phase and the actual experiment; for FPS: during startle habituation, CS habituation and fear acquisition training) led to exclusion of data from eight subjects for EMG analyses and 23 subjects for SCR analyses in the ASM study and four subjects for EMG analyses and data of eleven subjects for SCR analyses in the FPS study.

ASM: Significant effects were followed up via post-hoc t-tests to specify differences across categories. For valence ratings and EMG responses one-sided t-tests were performed as strong a priori assumptions exist for the direction of effect (i.e. positive>neutral>negative for valence ratings; negative>neutral>positive for EMG responses). For ratings of arousal as well as SCRs, one-sided t-tests were only performed for comparisons between emotional and neutral conditions whereas a two-sided t-test was performed between both emotional categories as no hypothesis exists for differences between negative and positive valences.

Functional magnetic resonance imaging (fMRI)

Data acquisition and processing

For both studies, MR data were acquired on a 3T MRI scanner [MAGNETOM Trio, Siemens, Erlangen, Germany) using a 12-channel head coil]. A high-resolution T1-weighted structural image (1x1x1mm) was acquired using a magnetization prepared rapid gradient echo sequence (MPRAGE).

ASM: Imaging parameters were specifically tailored to the brainstem and amygdala as our prime regions of interest: Twenty-five continuous axial slices (2 mm thick, no gap) were acquired using a T2*-sensitive gradient echo-planar imaging (EPI) sequence [repetition time (TR): 2.0 s; echo time (TE): 27 ms; flip angle: 70°; field of view (FOV): 232 x 232 mm, 2 x 2 mm in-plane resolution] in three sessions. TE was minimized using a parallel acquisition technique (generalized auto-calibrating partially parallel acquisitions, GRAPPA) with an acceleration factor of 2 and 24 reference lines. The slice package was adjusted to cover the lower border of the pons and the amygdala on the upper side. To avoid scanner drift artifacts over time, the experiment was divided into three scanning sessions in between which the scanner was re-adjusted without any interaction with the subject.

FPS: Imaging parameters were selected to cover the brainstem and simultaneously achieve near-complete coverage of the brain. 37 continuous axial slices (2 mm thick, 1 mm gap) were acquired using a T2*-sensitive gradient echo-planar imaging (EPI) sequence [TR: 3.0 s; TE: 26 ms; flip angle: 90°; FOV: 220 x 220 mm, 2 x 2 mm in-plane resolution]. TE was minimized using GRAPPA with an acceleration factor of 2 and 48 reference lines.

Data was processed within SPM12 (http://www.fil.ion.ucl.ac.uk/spm/) running on MATLAB2013a (The MathWorks, Natick, MA, USA). Initial fMRI preprocessing steps included discarding the first four volumes of each time series to account for T1 equilibrium effects, slice-time correction [for ASM (Study 1), using a HRF oversampling protocol], realignment and motion correction using the unwarp function implemented in SPM12.

Acoustic startle probes may present a potential source of task-related motion and therefore we used the "realign and unwarp" procedure. As such, unwarping corrects for potential residual/undetectable motion-related effects and has been demonstrated to successfully model distortion-by-movement related variance while leaving "true" activations intact (15). Following a reviewer's request to additionally include motion parameters, we have rerun all analyses with included motion parameters as nuisance regressors. Importantly, all statistically significant results remain statistically significant with some F-, T- and p-values for some analyses beeing larger while others are smaller when using suggested preprocessing pipeline. One result that was subthreshold (not) significant when using our initial preprocessing pipeline now is statistically significant (p_{uc}=0.001). Another exploratory result that we reported on an exploratory level (p_{uc} = 0.002) now just falls below our exploratory threshold of 0.005 (p_{uc} = 0.006) which is also true for the activation in the secondary PAG ROI in the ASM study (now p_{uc} = 0.002).

Following an additional request by a reviewer, we here provide mean and standard deviation of motion parameters for both studies (see **Table S2**).

Table S2. Motion parameters per study and per direction in mean [standard deviation] mm/degree.

Study	X	Y	Z	Pitch	Roll	Yaw
ASM (N=43)	0.084 [0.0577]	0.1201 [0.0789]	0.2078 [0.1461]	0.2460 [0.1736]	0.1272 [0.0880]	0.1133 [0.0755]
FPS (N=55)	0.1248 [0.0754]	0.1445 [0.0888]	0.3400 [0.1851]	0.3533 [0.2228]	0.2020 [0.1051]	0.175 [0.0954]

Note: Mean displacement represents deviation from the first image of the time-series. Mean was calculated as absolute values of displacements (i.e. negative values were transformed to positive values) to capture absolute value of displacement. Sign information was retained for calculation of standard deviation to capture full range of displacement in any direction.

In addition, prior to data acquisition of these studies, we have conducted extensive pilot testing of potentially occurring head motion and employed extensive and careful padding of participant's head in the MR-coil. Participants were instructed carefully with respect to the importance of laying still. In line with this previous work, manual inspection of motion parameters after realignment did not provide reason for concerns (e.g. motion parameters did not reveal translations exceeding two time the voxel size except for two participants in the FPS study which were primarily characterized by slow drifts). Most importantly, visual inspection

of the movement parameters did not suggest the occurrence of stimulus-correlated movement. Next, functional data was co-registered with the structural image. Two spatial normalization procedures of functional data were employed to account for standardized cortical imaging and the specific requirements for brainstem imaging, respectively. First, for analyses of the whole coverage of the FOV, data was normalized using DARTEL (16) and spatially smoothed with an isotropic Gaussian kernel (6mm FWHM) for analyses of whole-brain effects with specific focus on the amygdala as region of interest). Second, for improved brainstem spatial normalization, data was normalized using the SUIT toolbox as implemented for SPM (17) including up-sampling the data to a resolution of 1x1x1m. Since the target brainstem region of interest (i.e. the nucleus pontis caudalis, PnC) is a very small nucleus located within the pons, no spatial smoothing of functional data was applied during brainstem specific normalization (18). Note that all presented coordinates obtained from the brainstem specific analyses are in reference to the space defined by the SUIT toolbox and are in close alignment with the MNI space. During statistical estimation, further processing included temporal high-pass filtering (cut-off 128s) and correction for temporal auto-correlations using first-order autoregressive (AR1) modeling. Additionally, for brainstem specific statistical analyses, physiological noise correction was performed by adding 18 regressors of no interest using RETROICOR (19) which were estimated based on individual physiological data of cardiac (pulse curve recorded via pulsoxymeter) and respiratory data both acquired with a MR compatible monitoring system (Expression, InVivo, Gainesville, USA).

For illustrative purposes, parameter estimates of analyses were extracted using the rfxplot toolbox as implemented in SPM (20).

fMRI results are illustrated as statistical maps (t-values) overlaid on the CIT168 T1w 700µm template as provided by (21).

Data analyses of the primary startle pathway

For parametric analyses, the first-level model was extended by modulation of the habituation phase startle probe regressor with values of the time-dependent EMG habituation pattern that was observed across all participants (**Figure S2B**). In this approach trial-by-trial mean T-transformed EMG amplitudes were used in order to compensate for the limited number of trials and missing data. Note, this pattern is based on a mean responding of participants and does not include individual responses, as this analysis is based on only eight data points and missing data would hence reduce sample size and sensitivity within subjects. Estimated parameters for the modulated regressor are taken to the second-level by means of a one-sample t-test investigating a potential link between the neural activation within the regions of interest and the time-dependent EMG response pattern.

Data analyses of modulatory startle pathway

Valence-specific categorical analyses: To allow for the integration of the continuous EMG signal across sessions into of the fMRI data, data of all three sessions were concatenated. All regressors were convolved with a canonical HRF. Second-level analyses used SPM's flexible factorial model which permits correction for possible non-sphericity of the error term and takes a subject factor into account when analyzing differences in within-subject conditions. Within

this framework, first, a non-directional F-test was carried out to investigate differential neural activation to startle probes within the regions of interest (see below) between all three valence conditions (i.e., main effect: condition). Following up on these results, particular interest was directed at the *a priori* expected neural activation related to sub-cortical and brainstem responding to the startle probe during emotional conditions of negative-valence states (i.e. startle potentiation) as compared to positive-valence states (i.e. startle inhibition). Therefore, follow-up directional contrasts for neural responses towards the startle probe in negative>positive conditions were investigated. To explore the neural response to emotional pictures, an additional non-directional F-test based on a flexible factorial model, including estimated parameters for the emotional condition blocks, was calculated.

For the FPS paradigm, a GLM was set up for statistical first-level analysis. Regressors were constructed for CS onsets separated by CS-type (CS+/CS-) as well as startle presentation (no-startle/startle) during the CS habituation as well as the fear acquisition training phase, respectively. For CS habituation, these regressors served as regressors of no interest. Moreover, four additional regressors modeling the onsets of the habituation startle probes, inter-trial startle probes during CS habituation, inter-trial startle probes during fear acquisition training as well as for the USs were built. Ratings across all phases were modeled in one regressor as blocks for the entire duration of each rating block. All regressors were convolved with a canonical HRF function. Note, the startle probe regressor in the fMRI model shows inherent collinearity with the US regressor due to its close proximity in time. While the jittered startle probe onset as well as its presentation in only 66% of trials already reduces interpretation problems, this collinearity occurs only on the first (i.e., individual) level. Importantly, collinearity at the first level (as opposed to the second level) is not subject to estimation problems or increased risk of false positives but may result in highly variable parameter estimates and hence decreased sensitivity (22). *A priori* directional t-contrasts were calculated for the hypothesized effect of interest for increased startle probe onset reactivity during CS+ (threatening/stressful) as compared to CS- (safe/not stressful) conditions (i.e., CS+>CS-) within the fear acquisition training phase. Second-level analysis used SPM's one-sample t-test to test for significant differences across all individuals within the pre-defined regions of interest (see below).

EMG signal-integrative parametric analyses: First-level models designed for integrated eye-blink response data were similar to both models used in categorical analyses for ASM and FPS. However, for both studies, onsets for all startle probe regressors contained in one design matrix (i.e., startle probe onsets during conditions, startle probe habituation, and inter-trial-intervals) were condensed into one single regressor of interest. To assess the correlative relationship between neural and muscular activation, recorded raw EMG magnitudes were used as parametric modulator of the startle probe onset regressors. Raw EMG magnitude values were used because of the summary statistics approach and the centering procedure implemented for parametric modulators within SPM12. When blink responses were classified as missing values, these startle probe onsets were excluded from the startle regressor of interest and added to the design model as single regressor of no interest. To guarantee a stable parameter estimation and thereby a meaningful association between EMG magnitude values and neural activity, subjects having more than 33% missing values within the entire experimental phase were excluded from further second-level analyses. This led to reduced sample sizes for second-level analyses based

on integrated eye-blink data in the ASM ($N_{Categorical}$ = 43 vs. $N_{Integrated}$ = 29) as well as the FPS ($N_{Categorical}$ = 55 vs. $N_{Integrated}$ = 45) study. Second-level analyses were performed on the estimated parameters for the parametric modulator as calculated within the individual first-levels. A one-sample t-test was performed to find significant associations between neural and muscular activity.

Regions of interest and correction for multiple comparisons

For both studies, analyses focused on the centromedial amygdala (CMA) and PnC as main regions of interest while BNST (not covered in ASM) and PAG regions were investigated as secondary regions of interest for exploratory purposes.

Based on pre-defined masks for our regions of interest [CMA, centromedial amygdala ((23) within the Jülich SPM Anatomy Toolbox (v2.1.) (24)) which is based on maximum probability maps rather than individual thresholding (25); BNST thresholded at 50% (26) and PAG (27)], multiple comparisons were controlled for by using a small-volume correction (SVC) approach using family-wise error correction (FWE_{SVC}<0.05, cluster-forming threshold at 0.001). Visualization of statistical maps are displayed as thresholded but uncorrected effects in the main text figures. For comparison with corrected effects, see **Figure S1**.

All masks used for small volume corrections are uploaded to Neurovault and can be inspected at https://neurovault.org/collections/4469/

Given that this is the first high-resolution fMRI study targeting the PnC and thus no coordinates as derived from fMRI exist, the brainstem-specific analyses targeting the PnC are reported on a conventional uncorrected threshold of p<0.001. The PnC location was identified by converging information based on anatomical correspondence between structural MRI landmarks and definitions as provided by Duvernoy's Atlas of the Human Brain Stem and Cerebellum (Naidich et al., 2009, **Figure 1A**). This identified location is additionally supported by its location in reference to a just recently available MRI mask of the nucleus reticularis pontis oralis (PNO, (27), **Figure 1A**) while availability of a MRI mask of the PnC is still lacking.

Supplementary Results

Eye-blink responses in the habituation phase of the ASM study followed a typical trial-by-trial habituation pattern with decreasing response magnitudes over trials. Based on this, parametric modulation analyses revealed two clusters of the CMA (see **Table S3 '**Habituation pattern'**, Figure S2A**) and a cluster in the proximity to the PnC region (**Table S3**) mapping onto the habituation pattern of eye-blink response strength (**Figure S2B/C**).

Table S3. Statistics for activation in the PnC region to startle probe onset during startle probe habituation (unmodulated responding) as well as activation mirroring the EMG habituation pattern in the PnC and secondary regions of interest (CMA, PAG).

PnC		**$p_{uncorrected}$**	**k**	**T**	**X**	**Y**	**Z**
Unmodulated responding							
		<0.001	3	3.47	2	-35	-36
Habituation pattern							
		<0.001	2	3.63	6	-32	-35
Secondary ROIs[a]		**$p_{FWE(SVC)}$**	**$k_{(SVC)}$**	**T**	**X**	**Y**	**Z**
CentromedialAmygdala							
Unmodulated responding							
	left	<0.001	80	5.23	-21	-4	-14
	right	<0.001	24	4.90	24	0	-15
		<0.001	33	4.89	20	-9	-14
		<0.001		4.80	26	-9	-12
		<0.001		4.65	22	-6	-12
Habituation pattern							
	left	0.008	3[b]	3.75	20	-9	-14[c]
		0.013	1	3.55	20	-4	-15
PAG							
Unmodulated responding		0.004	13	4.16	-3	-32	-10
			9	3.58	4	-32	-10

[a] Note, because of the restricted FOV BNST activation was not assessed.

[b] the CMA clusters are part of a larger amygdala cluster with an unrestricted cluster-extend of 120 voxels which lies very specifically within the amygdala boundaries, it might be cautioned to attribute this relationship solely to the area of the CMA. Other areas of the amygdala might also be involved in this habituation related relationship and warrant further investigations.

[c] Note, this is the same peak voxel coordinate as in unmodulated responding.

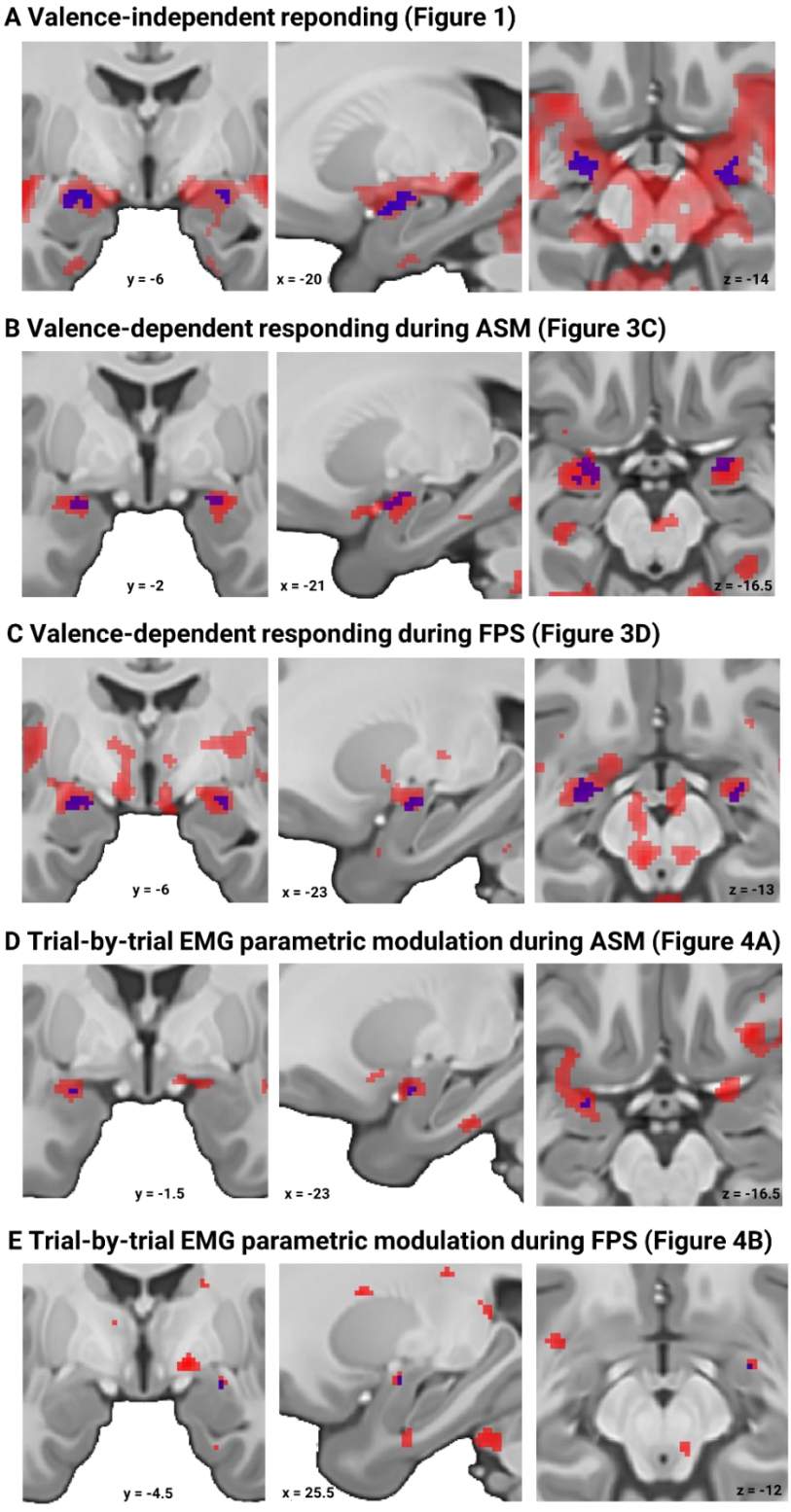

Figure S1: Visualization of the CMA results after correction for multiple comparisons using small volume correction (blue) overlaid on the thresholded but uncorrected effects (red) as displayed for the main analyses presented in the main text. The corresponding figure number in the main manuscript is provided as a reference in brackets. Overlap shows strong correspondence between the proposed region of interest (i.e. CMA) and the unmasked cluster extends for all analyses.

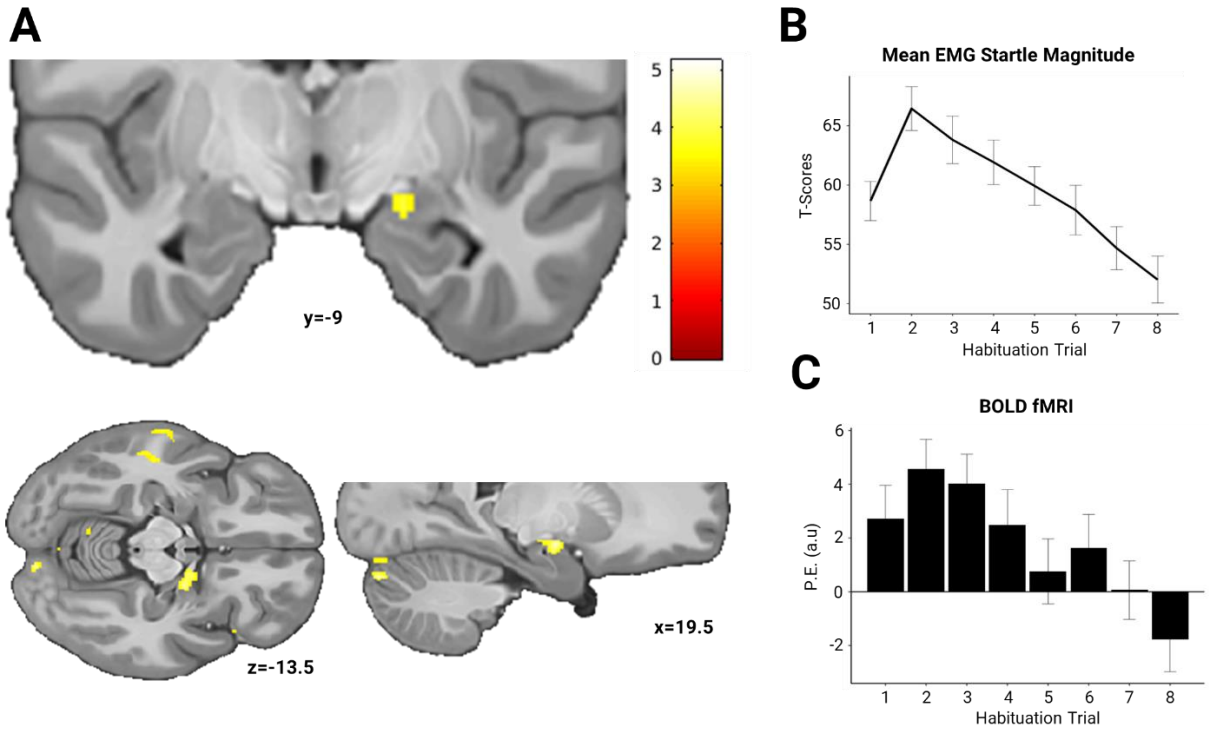

Figure S2. (A) Functional mapping of amygdala activation onto the mean trial-by-trial EMG habituation pattern as shown in (B). (C) Extracted estimated parameters from session 1 for the peak amygdala voxel in (A) mirroring the EMG pattern in (B). Display threshold at p<0.001. P.E.(a.u.): parameter estimates (arbitrary units).

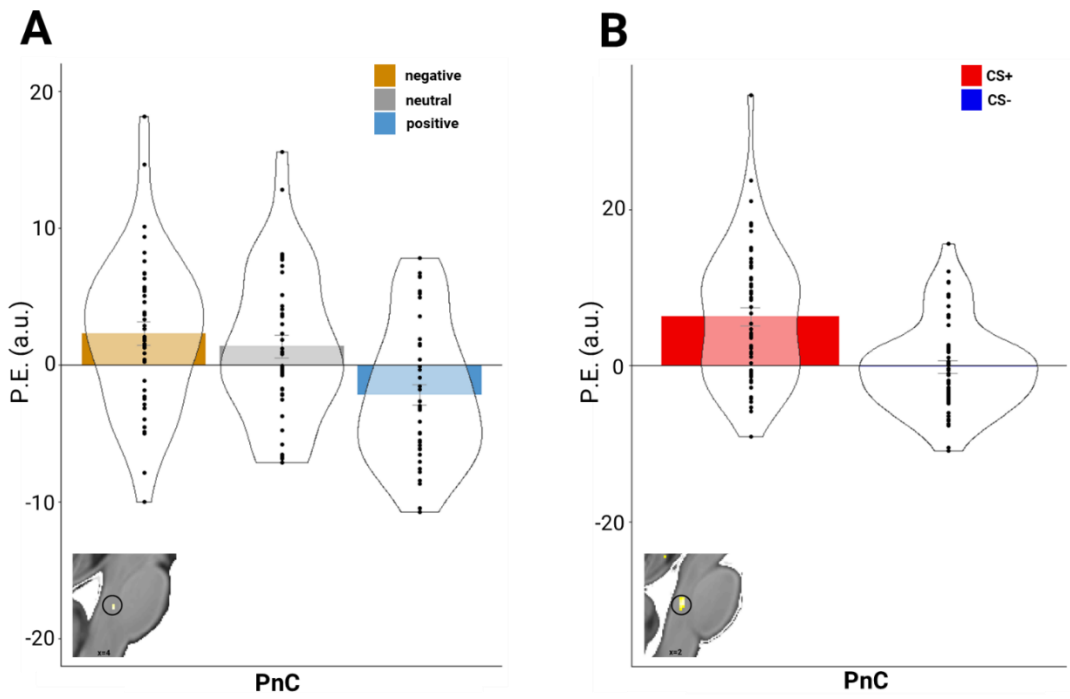

Figure S3. Individual data points and density distribution for peak voxel parameter estimates of the PnC area for valence-dependent neural activation evoked by startle probes (as presented in **Figure 3** within the main text) for (A) during ASM: negative>positive and (B) FPS: CS+>CS.

Supplementary Discussion

Physiological habituation processes are known to affect activation in the PnC region in the primary acoustic startle reflex in rodents (29). In line, we observe a characteristic habituation pattern in the PnC with activation strength mirroring EMG magnitude (**Figure S2)**. A similar habituation pattern was observed in the CMA. Importantly, the initial habituation phase may be perceived as unpleasant and the habituation-like pattern may in fact reflect adaptation to the aversive situation. Consequently, adaptation to aversiveness may be considered as a type of affect modulation. It can thus be speculated that the response pattern as observed in the PnC may at least partially reflect affective modulation processes. As such these findings can be taken to support our primary hypotheses presented within the main text.

Supplementary References

1. Bradley MM, Codispoti M, Lang PJ (2006): A multi-process account of startle modulation during affective perception. *Psychophysiology*. 43: 486–497.

2. Bradley MM, Cuthbert BN, Lang PJ (1993): Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology*. 30: 541–545.

3. Lang PJ, Bradley M, Cuthbert B (2008): International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. *Tech Rep*. University of Florida, Gainesville, FL.

4. Wessa M, Kanske P, Neumeister P, Bode K, Heissler J, Schönfelder S (2010): EmoPics: Subjektive und psychophysiologische Evaluationen neuen Bildmaterials für die klinisch-bio-psychologische Forschung. *Zeitschrift für Klin Psychol und Psychother*. 1/11.

5. Brainard DH (1997): The Psychophysics Toolbox. *Spat Vis*. 10: 433–6.

6. Sjouwerman R, Niehaus J, Kuhn M, Lonsdorf TB (2016): Don't startle me—Interference of startle probe presentations and intermittent ratings with fear acquisition. *Psychophysiology*. 53: 1889–1899.

7. Bradley MM, Lang PJ (1994): Measuring emotion: The self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 25: 49–59.

8. Mandelkow H, Halder P, Boesiger P, Brandeis D (2006): Synchronization facilitates removal of MRI artefacts from concurrent EEG recordings and increases usable bandwidth. *Neuroimage*. 32: 1120–1126.

9. Lindner K, Neubert J, Pfannmöller J, Lotze M, Hamm AO, Wendt J (2015): Fear-potentiated startle processing in humans: Parallel fMRI and orbicularis EMG assessment during cue conditioning and extinction. *Int J Psychophysiol*. 98: 535–545.

10. Blumenthal TD, Cuthbert BN, Filion DL, Hackley S, Lipp O V, van Boxtel A (2005): Committee report: Guidelines for human startle eyeblink electromyographic studies.

Psychophysiology. 42: 1–15.

11. Haaker J, Gaburro S, Sah A, Gartmann N, Lonsdorf TB, Meier K, *et al.* (2013): Single dose of L-dopa makes extinction memories context-independent and prevents the return of fear. *Proc Natl Acad Sci U S A*. 110: E2428-36.

12. Boucsein W, Fowles DC, Grimnes S, Ben-Shakhar G, Roth WT, Dawson ME, Filion DL (2012): Publication recommendations for electrodermal measurements. *Psychophysiology*. 49: 1017–1034.

13. Venables P, Christie M (1980): Electrodermal activity. *Tech Psychophysiol*. Chichester, UK: Wiley, pp 3–67.

14. Lykken DT, Venables PH (1971): Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*. 8: 656–672.

15. Andersson JLR, Hutton C, Ashburner J, Turner R, Friston K (2001): Modeling Geometric Deformations in EPI Time Series. *Neuroimage*. 13: 903–919.

16. Ashburner J (2007): A fast diffeomorphic image registration algorithm. *Neuroimage*. 38: 95–113.

17. Diedrichsen J (2006): A spatially unbiased atlas template of the human cerebellum. *Neuroimage*. 33: 127–138.

18. Beissner F (2015): Functional MRI of the Brainstem: Common Problems and their Solutions. *Clin Neuroradiol*. 25: 251–257.

19. Glover GH, Li TQ, Ress D (2000): Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med*. 44: 162–167.

20. Gläscher J (2009): Visualization of group inference data in functional neuroimaging. *Neuroinformatics*. 7: 73–82.

21. Pauli WM, Nili AN, Tyszka JM (2018): A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. *Sci Data*. 5: 180063.

22. Mumford JA, Poline JB, Poldrack RA (2015): Orthogonalization of regressors in fMRI models. *PLoS One*. 10.

23. Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ, *et al.* (2005): Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: Intersubject variability and probability maps. *Anat Embryol (Berl)*. 210: 343–52.

24. Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*. 1: 1325–35.

25. Eickhoff SB, Heim S, Zilles K, Amunts K (2006): Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *Neuroimage*. 32: 570–

582.

26. Torrisi S, O'Connell K, Davis A, Reynolds R, Balderston N, Fudge JL, *et al.* (2015): Resting state connectivity of the bed nucleus of the stria terminalis at ultra-high field. *Hum Brain Mapp*. 36: 4076–88.

27. Edlow BL, Takahashi E, Wu O, Benner T, Dai G, Bu L, *et al.* (2012): Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. *J Neuropathol Exp Neurol*. 71: 531–46.

28. Naidich TP, Duvernoy HM, Delman BN, Sorensen AG, Kollias SS, Haacke EM (2009): *Duvernoy's Atlas of the Human Brain Stem and Cerebellum: High-Field MRI, Surface Anatomy, Internal Structure, Vascularization and 3 D Sectional Anatomy*. Springer Science & Business Media.

29. Koch M (1999): The neurobiology of startle. *Prog Neurobiol*. 59: 107–128.