Expectation of mood improvement entails resource-dependent downregulation of negative inputs

Jonas Rauh[1], Arasch Mostauli[1], Matthias Gamer[2], Christian Büchel[1], Winfried Rief[3] & Stefanie Brassen[1*]

[1] Department of Systems Neuroscience, University Medical Centre Hamburg-Eppendorf, Martinistrasse 52, D-20246, Hamburg, Germany

[2] Department of Psychology, University of Würzburg, D-97070 Würzburg, Germany

[3] Department of Clinical Psychology and Psychotherapy, Philipps-University Marburg, D-35037 Marburg, Germany

[*]Correspondence: sbrassen@uke.de

Abstract

Clinical trials of antidepressants show improvements in placebo groups of about 80% compared to the real treatment arm. The mechanism underlying this clinically important effect has been linked to goal-directed attentional control. Here, we investigated in detail how attentional resources and cognitive control influence the effects of positive expectations on emotional processing. Forty-nine healthy volunteers participated in a cross-over fMRI study, in which expectancy of mood improvement and emotional processing was induced by an alleged oxytocin nasal spray and verbal instruction. In the MR scanner, a spatial cueing paradigm manipulated attentional load on emotional face distractors. Behavioral findings showed placebo treatment to improve mood and to reduce distractibility by fearful over happy faces, specifically when more attentional resources were available to process faces. This aligned with neural changes in activation and functional coupling in lateral prefrontal-limbic networks indicating an expectation-induced top-down regulation of aversive inputs. In addition, effects in behavior and prefrontal-parietal networks directly correlated with individual cognitive control. Our data show that expectations recruit top-down attentional selection and that this influence critically depends on individual control ability and attentional context. These findings may be particularly relevant in patients with major depressive disorder, who often demonstrate a negativity bias and reduced cognitive control capacity.

Introduction

There is an ongoing debate about the influence of patient expectations on the effectiveness of antidepressant treatment. Studies show that placebo groups can achieve about 80% of the benefits observed in groups receiving antidepressants. Patient expectations are thought to be a key mediator of these clinical effects (Kirsch, 2019; Leuchter et al., 2014; Rief et al., 2009; Rutherford et al., 2017; Stone et al., 2022; Thase et al., 2011). Insights into the mechanisms of placebo effects on mood state and emotion processing have the potential to improve treatment strategies for depression. However, to identify mechanisms underlying placebo responses, it is necessary to contrast the placebo condition to a control condition without expectation induction.

There is solid evidence that placebo treatments can substantially modulate affect (Crichton et al., 2014; Geers et al., 2021). For instance, positive expectations can reduce experimentally induced sadness (Anna et al., 2019; Haas et al., 2020; Schaefer et al., 2023), anxiety or fear (Meyer et al., 2015, 2019; Petrovic et al., 2005; Shafir et al., 2023; Zhang et al., 2011) and social stress (Abrams et al., 2001; Balodis et al., 2011). Moreover, placebo manipulations can positively influence baseline mood states (Baker et al., 2022; Zhang et al., 2011). Interestingly, such effects are still observable when the placebo manipulation is not deceptive as in open-label placebo studies (Hahn et al., 2022; Schienle et al., 2022; Sezer et al., 2022).

Imaging studies of placebo effects in the affective system often report an involvement of the lateral prefrontal cortex (PFC) (Peciña et al., 2018; Petrovic et al., 2005), the ventromedial PFC (vmPFC) (Ellingsen et al., 2013), and the anterior cingulate cortex (ACC) (Petrovic et al., 2005; Schaefer et al., 2023; Zhang et al., 2011). Prefrontal engagement has been attributed to the construction and maintenance of instructed beliefs (Geuter et al., 2017; Koban et al., 2017). In particular, the lateral PFC seems to be pivotal for building and updating of treatment expectations, and the signaling of instructed state representations to bias perceptual processing in downstream networks (Koban et al., 2017; Wager et al., 2004).

The consistent finding of an involvement of prefrontal regulation in expectation effects suggests a critical role of cognitive functioning and goal-directed attentional control (Corbetta & Shulman, 2002). In a behavioral study using cognitive distraction, however, placebo analgesia occurred

independently from an active redirection of attention from pain, which suggests a rather limited role of attentional resources (Buhle et al., 2012). On the other hand, reduced executive functioning and weaker prefrontal resting-state EEG connectivity have been associated with disrupted placebo responses in Alzheimer's disease patients (Benedetti et al., 2006) and cognitive reappraisal ability mediated via the dorsolateral PFC has been positively related to placebo analgesia in young adults (van der Meulen et al., 2017). Research on the role of cognitive resources in expectation effects has mainly focused on the pain system, while far less is known about such interactions in the affective system. However, understanding the role of cognitive control in expectation effects related to affective states seems particularly important for extrapolating findings to clinical conditions such as major depression, in which prefrontal and executive functions are often impaired (Grahek et al., 2019; Millan et al., 2012; Pizzagalli & Roberts, 2022; Snyder, 2013).

A promising approach to illuminate the role of cognitive control for expectation effects is to directly combine the manipulation of cognitive resources and expectations with functional neuroimaging of prefrontal networks during emotional processing. In the present study, 49 healthy volunteers, characterized in their general cognitive control ability, participated in a randomized, cross-over fMRI study that was combined with a controlled induction of positive expectations (Baker et al., 2022). During fMRI, participants performed a spatial cueing paradigm that manipulated attentional load on emotional face distractors and thus attentional resources to process emotional stimuli in a goal directed manner (Brassen et al., 2010, 2011). Specifically, in this paradigm an irrelevant, centered background face receives low attentional resources when in the task an arrow shifts the attentional focus to either the right or left side of the display screen. In contrast, when no directional cue is presented, the attentional focus is not shifted, and more attentional resources remain for (controlled) face processing. We hypothesized that positive expectations influence emotional processing of faces congruently when more attentional resources are available and that these effects are modulated by prefrontal networks and individuals' overall ability to control their attention.

Results

Study overview. Forty-nine healthy young adults (age: 26.92 ± 4.23 years; 30 women) participated in three study days. On the first day (day 0), participants were screened for inclusion criteria and their general ability to exert top-down control over salient distraction was measured with a visual search task (Singleton-task; Theeuwes, 1991, 1992). On the following days (day 1 and 2), all participants underwent fMRI scanning, separated by at least 1 week. Participants received either intranasal alleged oxytocin (in the following: condition 'placebo') or saline (in the following: condition 'control') in a randomized cross-over design (Figure 1). Expectations about positive effects of the oxytocin application were induced by verbal instructions (e.g., "Oxytocin can inhibit the perception of aversive stimuli such as fearful faces") and a standardized video documentary. Treatment expectations and experiences were assessed through standardized questionnaires (Rief et al., 2017). Mood ratings were acquired via visual analogue scales (VAS). In the scanner, participants had to respond by button press to a left- or right-sided dot-target, which was preceded by a spatially informative (directed), or uninformative (non-directed) cue. Cues were superimposed on emotional faces or scrambled pictures, serving as distracters (Figure 1). As demonstrated previously (Brassen et al., 2010, 2011), during the non-directed cue condition, more attentional resources are available to process centered distractors. Participants were instructed to respond as quickly and as accurately as possible to the side of target presentation and that faces are irrelevant for the experiment. The event-related paradigm was divided into three blocks.

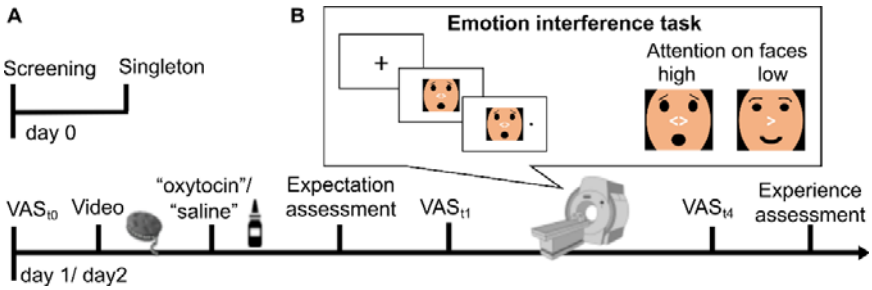

Figure 1. Study design and paradigm. (A) Timeline. **(B)** Emotion interference task using a stylized example stimuli set to illustrate trials with high (undirected cue) and low (directed cue) attentional load on face distractors. Stimuli from the Karolinska Directed Emotional Faces Database (KDEF) were used for the paradigm, but due to preprint policy, only stylized stimuli were shown here for illustration. VAS = visual analogue scale.

Positive expectation induction enhances mood state and post-treatment valuation

Participants rated their expectation of positive treatment effects (i.e., improvement of mood) as significantly higher in the placebo treatment than in the control session ($T_{(48)} = 7.90$, $p < .001$, $d = 1.13$, Figure 2A), validating our experimental manipulation.

On each fMRI study day, mood state was assessed via a visual analogue scale (VAS) at participants' arrival (baseline VAS$_{t0}$), after alleged 'oxytocin' or 'saline' application, and after each of the three scanning blocks (VAS$_{t1-4}$). Baseline mood state did not differ between sessions ($p = 0.12$). rmANOVA on baseline corrected VAS values with the factors 'condition' (control/placebo) and 'time' (t1-t4) demonstrated a strong condition effect ($F_{(1,48)} = 8.61$, $p = .005$, $\eta^2 = 0.15$, Figure 2B) that was due to higher mood ratings for the placebo (M = 3.39 ± 5.25) compared to the control session (M = -14.58 ± 5.26). In addition, there was a main effect of time ($F_{(3,46)} = 8.16$, $p < .001$, $\eta^2 = 0.35$) mainly showing a relative mood decrease towards the end of the scanning sessions, probably indicating habituation and fatigue effects.

Because previous treatment experience is an important prior for future treatment success, participants' overall treatment experience was also assessed at the end of each session. Participants reported a stronger experience of mood increase at the end of the placebo treatment session as compared to the control session ($T_{(48)} = 5.14$, $p < .001$, $d = 0.73$, Figure 2C). This effect was directly correlated with positive effects of placebo on mood state ($r = .31$, $p = .029$).

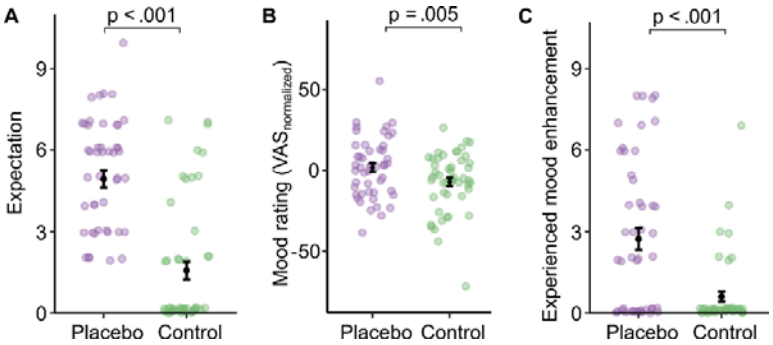

Figure 2. Placebo effects on expectation, mood, and experience ratings. Individual data with group means and s.e.m.

Placebo-related reduced distractibility by fearful faces is mediated by attentional context and cognitive control

On both study days, participants showed high accuracy in task performance indicated by a very low number of incorrect responses (control: 3.0% ± 7.6%; placebo: 1.3% ± 2.0%). Consistent with our previous findings (Brassen et al., 2010, 2011), there was a strong manipulation of attention by the different cues as reflected by a significantly increased reaction time under the neutral, non-directing cue (NC) compared to the valid, directing cue (DC) condition across all trials on the control ($T_{(48)}$ = 12.09; $p < .001$, $d = 1.73$) and the placebo session ($T_{(48)}$ = 10.66; $p < .001$, $d = 1.52$), which did not differ significantly from each other ($F_{(1,48)}$ = 0.04; $p > .85$). Eye movement data were recorded to control for fixation during the cueing period and data of 27 participants could be analyzed. In line with our previous findings (Brassen et al., 2010, 2011), participants were well able to maintain fixation, even when directional cues were presented. Horizontal gaze shifts were well below 0.5° of visual angle and an ANOVA on gaze positions did not yield any significant effect (all $p > .05$; see Supplementary information).

After-scan arousal ratings on the control day showed a significant effect of emotion ($F_{(3,46)}$ = 37.45; $p < .001$, $\eta^2 = 0.71$). Post-hoc analyses revealed these effects to be caused by the following ranking of arousal ratings: happy ≈ fearful > sad > neutral. There was no significant difference to post-scan ratings from the placebo session ($F_{(3,46)}$ = 1.28; $p > .29$). As expected, post-scan valence ratings strongly matched pre-defined emotional expressions with an overlap of 94% on each of both sessions.

As we were mainly interested in expectation effects on the interference by specific emotional faces and how this may be modulated by attentional resources, we first identified baseline interference by different emotional compared to neutral faces for the different attentional conditions. To this end, for each emotion separately, reaction times for neutral and emotional trials from the control session were entered into an rmANOVA with the within-subject factors emotion (neutral vs. fearful or happy or sad) and attention to faces (high vs. low). A significant main effect of emotion was only obtained for fearful versus neutral faces ($F_{(48,1)}$ = 12.64, $p < .001$, $\eta^2 = 0.21$) which was due to significantly slower RTs in fearful trials. There was no main effect of emotion for happy or sad faces (all $p > .16$) nor were there any emotion × attention interactions (all $p > 0.31$, Supplementary information).

Because a bias for fearful faces could also be influenced by intensity differences from neutral trials, we directly compared fearful to happy trials in another rmANOVA. Again, there was as a significant main effect of emotion ($F_{(1,48)}$ = 9.97, $p = .003$, $\eta^2 = .17$) but now there was also a significant emotion × attention interaction ($F_{(1,48)}$ = 4.69; $p = .035$, $\eta^2 = .09$). Results were driven by generally slower RTs in fearful compared to happy trials which was particularly pronounced in the high attention to faces condition ($T_{(48)}$ = 3.47, $p < .001$, $d = .50$). We then tested whether this bias for fearful faces changed for the placebo treatment using paired T-Tests (two-tailed). Results showed that the bias (i.e., slower RTs for fearful compared to happy trials) in the high attention condition decreased significantly on placebo ($T_{(48)}$ = 2.23, $p = .031$, $d = 0.32$) while there was no significant change in the low attention condition ($P > 0.85$, Figure 3). There were no placebo effects in control trials (scrambled faces across and within attentional conditions, all $p > 0.81$), indicating that placebo had no general effect on task performance.

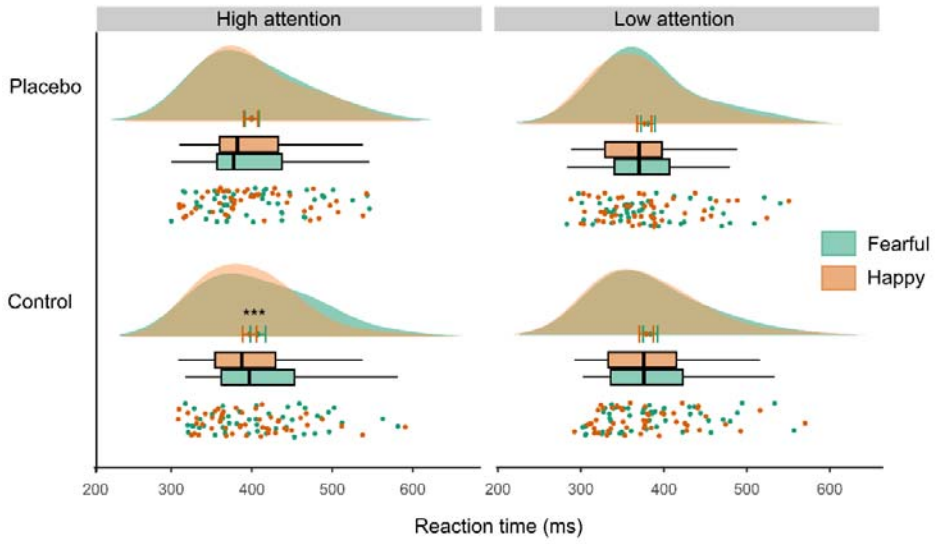

Figure 3. Behavioral results in the emotional interference paradigm. Distraction by fearful compared with happy faces at control (***p < .001) was significantly reduced at placebo (interaction: p < .04). Findings were restricted to the high attention to faces condition. Graph depicts individual data, boxplots, mean with s.e.m., and data distribution.

Finally, we tested for a possible modulation of significant placebo effects by individuals' expectation ratings and cognitive control ability. For this purpose, two ANCOVAs were performed including session (control vs. placebo), emotion (fearful vs. happy) and attention to faces (high vs. low) as well as the respective covariates (expectation ratings / Singleton score). While there was no significant interaction with subjective expectation ratings, results yielded a significant relationship between individual singleton scores and the session by attention interaction ($F_{(1,47)} = 10.09$, $p = .003$, $\eta^2 = 0.18$). Post-hoc correlations showed that this interaction was mainly due to a specific impact of cognitive control on placebo effects in the high attention as compared to the low attention condition (Fisher's $Z = 2.91$, $p = 0.002$), and this effect was restricted to highly attended fearful trials ($control_{fearful_nc} > placebo_{fearful_nc}$ × Singleton: $r = -.41$, $p = 0.004$, Figure 4). That is, individuals with higher general cognitive control ability showed a specifically stronger bias reduction for fearful faces.

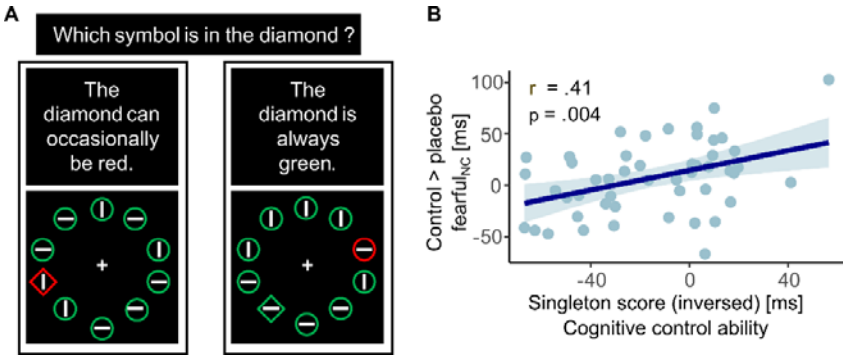

Figure 4. Correlation between behavioral findings and cognitive control ability. (A) Illustration of the two singleton-task conditions (for calculation of the singleton score, see Methods section) **(B)** Placebo effects on distractibility by fearful faces correlated with cognitive control ability.

Neural responses in emotion and attention networks to fearful faces decrease at placebo

First confirmatory analyses replicated our previous data on the emotional interference paradigm by demonstrating increased activation in primary and higher order visual processing networks in response to non-directed as compared to directed cues, validating the attentional manipulation in the task. Moreover, fMRI-data demonstrated an increased activity in a visual emotional network including the extrastriate visual cortex and bilateral amygdala in response to face stimuli as compared to scrambled pictures (Supplementary information). Including all trials, these patterns did not differ between sessions (control vs. placebo).

In our targeted analyses, we then concentrated on fearful and happy face trials, corresponding to the main behavioral results. To this end, we constructed an analysis of variance for fearful versus happy faces, including the contrast images of high vs. low attention to emotional faces from both sessions (control vs. placebo), resulting in a 2 × 2 × 2 factorial design. In accordance with the behavioral finding of an increased distractibility by highly attended fearful compared to happy faces on the control day, increased activity in response to fearful faces was observed in cortical midline structures, including the superior medial PFC, the middle cingulate cortex and the bilateral precuneus. In addition, increased activity was observed in the left middle temporal gyrus as well as in the superior frontal and superior parietal gyrus (Supplementary information, Table S1).

As one of our key analyses, we then addressed the question whether there was a placebo-related attenuation in the specific response of neural networks to highly attended fearful face stimuli. In accordance with the behavioral data, the activity in several areas that had been increased on the control day, significantly decreased in the placebo compared to the control condition (control > oxytocin: $fearful_{nc}$ > $happy_{nc}$), including the middle and anterior CC, precuneus, and left superior frontal gyrus (Figure 5B). Findings were restricted to the high attention (NC) condition, while no activation differences appeared for fearful versus happy face trials in the low attention condition (DC) on the control day or as a difference to the placebo condition (Table S1). Thus, in accordance with the behavioral data, findings indicate a specific placebo-related suppression of networks engaged in emotional and interference processing in response to highly attended fearful faces.

Prefrontal modulation of the placebo effect

The main focus of the present study was placed on the modulatory mechanisms involved in the affective placebo response with a special emphasis on prefrontal networks (Figure 5A) and cognitive

control ability. Here, the right anterior-lateral PFC was significantly activated, specifically in response to highly attended fearful faces, when placebo was compared with the control condition (placebo>control: $fearful_{nc} > happy_{nc}$, Figure 5C, Supplementary Table S1).

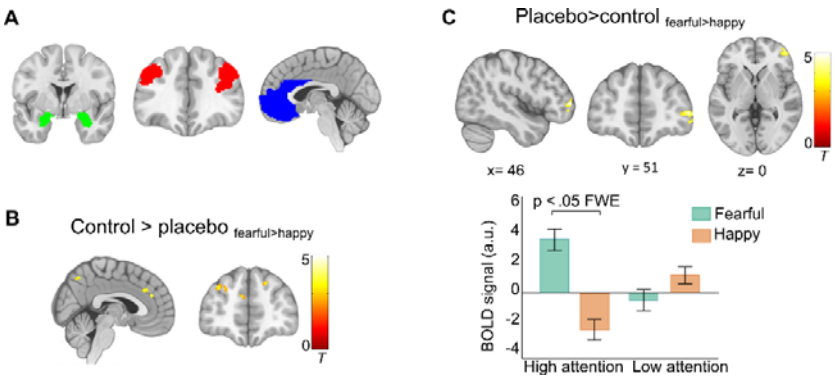

Figure 5. Neural placebo effects. (A) Predefined ROIs in the amygdala (green), dlPFC (red), and the vmPFC/ACC (blue) overlaid on the MNI structural template. **(B)** Placebo-induced signal decrease in response to highly attended fearful faces in the middle frontal and cingulate gyrus, the medial superior PFC and precuneus. **(C)** Placebo-induced signal increase in the right middle frontal gyrus. The bar graph shows group means and s.e.m. of mean parameter estimates (placebo > control) extracted from individual spheres within this activation cluster. All results are $p < .05$ FWE corrected. Activations are overlaid on a standardized MNI template (display threshold $p < .001$ uncorrected).

We then used psychophysiological interaction (PPI) analysis to assess whether the individual functional coupling of this prefrontal region changed under placebo compared with the control condition during the processing of highly attended fearful faces. Results showed a stronger negative coupling between the lateral PFC and the ACC/vmPFC as well as with the bilateral amygdala, indicating a placebo-induced increased prefrontal top-down regulation of limbic regions involved in emotional valuation (Figure 6A, Supplementary Table S2).

Placebo-induced prefrontal-parietal activity correlates with individual ability to control attention

Finally, to test for neural correlates of the observed relationship between placebo effects and individuals' attentional control ability we conducted a regression analysis including the contrast images $fearful_{placebo\ vs.\ control}$ from the high attention condition and individual singleton scores as a covariate. Individual attentional control ability was significantly correlated with activity in a prefrontal-parietal network for placebo compared to control. Specifically, significant clusters were observed in the right DLPFC and the bilateral superior parietal lobe (SPL, Figure 6B and Supplementary Table S3). Thus, individuals with better ability to control attention as observed in the Singleton task showed a stronger placebo-related increase in lateral prefrontal-parietal networks typically involved in attentional top-down control when being confronted with fearful face distractors.

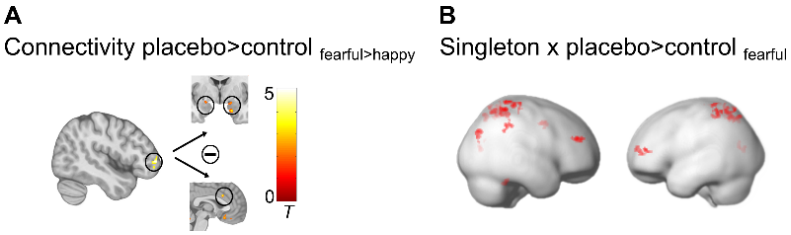

Figure 6. PPI and regression results. (A) Placebo-induced increased negative coupling of the right middle frontal gyrus with the bilateral amygdala and with the dorsal part of the ACC. **(B)** Cognitive control ability (Singleton) correlated positively with placebo-induced increased signals in a prefrontal-parietal network. All results are $P < 0.05$ FWE corrected. Activations are overlaid on standardized structural and rendered templates (display threshold $P < 0.005$ uncorrected)

Discussion

The current data demonstrate that verbally induced positive expectations not only improve subjective mood but also activate top-down regulation of irrelevant aversive stimuli in a context with high attentional resources for stimulus processing. Beneficial placebo effects were reflected in the down-regulation of networks involved in emotional valuation and cognitive interference in response to distracting fearful faces. Placebo-related changes were thereby mediated by increased engagement of the lateral PFC and a network of prefrontal-parietal regions that was directly associated with individual differences in the general ability to exert cognitive control. Changes in functional coupling of the lateral PFC with the vmPFC and the amygdala in the placebo condition further emphasize the importance of resource-dependent top-down regulation in verbally instructed placebo effects on affective processing. This is of particular relevance regarding opportunities and limitations of treatment expectations in clinical samples.

The placebo-related decrease in perception of aversive stimuli is consistent with previous data on placebo analgesia and affective expectation manipulation. Many studies used a combination of verbal instructions and conditioning to maximize expectation effects (e.g., Colloca 2010; Tinnermann 2017; Petrovic 2009). Here, we could show that such a decrease can also be achieved by verbal instruction alone. At the neurophysiological level, attenuated distraction by fearful face distractors was paralleled by an active suppression of neural networks processing emotional reactivity and interference (Etkin et al., 2015; Song et al., 2017) as well as visuospatial orienting (Corbetta et al., 2008), including the anterior and dorsal cingulate cortex, the superior prefrontal cortex, and the precuneus. Reduced BOLD signals in these regions indicate reduced interference due to less attentional distraction.

Verbally instructed placebo effects have been attributed to changes in expectations and appraisal processes which in turn may be implemented in prefrontal brain networks and their modulation of subcortical, limbic systems (Benedetti, 2014; Koban et al., 2017; Schwarz et al., 2016). Our data confirm these hypotheses, similar to previous studies on neural expectation effects in the affective system (Ellingsen et al., 2013; Petrovic et al., 2005). Specifically, reduced processing of fearful faces was associated with enhanced activity in regions of the right lateral PFC. In addition, connectivity analyses demonstrated an increased negative coupling of the right lateral PFC with the amygdala and the vmPFC for placebo treatment, suggesting increased inhibition of affective valuation processes by higher order cognitive control. The lateral PFC is pivotal in cognitively demanding explicit strategies and cognitive flexibility (Friedman & Robbins, 2022). In the context of expectation effects, its engagement has been associated with the top-down biasing of affective key nodes in the service of

the instructed state (Koban et al., 2017). This may be modulated through the vmPFC that is thought to integrate information about current goals from the lateral PFC with subcortical information, e.g., from the amygdala, about the relevance of stimuli to affective goals (Geuter et al., 2017; Haber & Knutson, 2010; Koban et al., 2017; Lapate et al., 2022). Thus, our data suggest that increased activity in the lateral PFC signals instructed beliefs about positive effects of oxytocin and biases appraisal processes in downstream networks accordingly through the inhibition of attention and value signals in response to aversive stimuli.

We focused on the role of attentional and cognitive resources in verbally instructed expectation effects in the affective system. Despite striking evidence for an involvement of prefrontal-limbic top-down regulation in placebo effects across systems (Atlas & Wager, 2014; Geuter et al., 2017; Koban et al., 2017; Zunhammer et al., 2021), results regarding the direct modulation through cognitive processes is rather heterogeneous. For example, previous data suggested that placebo analgesia does not rely on an active redirection of attention and is not reduced by cognitive distraction (Buhle et al., 2012). On the other hand, individual differences in cognitive reappraisal ability were directly related to down-modulating signals in the dorsolateral PFC during pain inhibition in placebo analgesia (van der Meulen et al., 2017). To our knowledge, no study so far investigated the role of cognitive resources in affective placebo effects, even though treatment expectation is crucial in mood disorders in which cognitive function can be attenuated. Moreover, while it seems obvious that changes in the selective perception of expectation-related stimuli depend on the specific ability to control attention, the role of this specific cognitive mechanism is unclear.

Replicating previous data (Brassen et al., 2010, 2011), our paradigm was able to modulate attentional resources available to process emotional face distractors, as evident by slower reaction times and pronounced neural face processing in the high compared with the low attention condition. Expectation of positive oxytocin effects specifically lowered distraction by fearful compared to happy faces, mirrored by aforementioned changes in appraisal and attentional control networks. Effects were restricted to neutral cue trials, where distraction was most pronounced, but also more resources were available for a goal-directed allocation of attention, consistent with instructed expectations (e.g., "Oxytocin can inhibit the perception of aversive stimuli such as fearful faces"). No placebo-related changes were observed in the low attention condition suggesting a resource dependency of this effect. One could also have expected a placebo-related increased distraction by happy faces consistent with the instruction that oxytocin enhances positive stimulus perception. Indeed, we found an induced positivity bias in face perception in a previous study on positive expectation effects in emotion detection (Baker et al., 2022). The fact that we did not see such a clear effect in the present study might be due to the different nature of the task. Here, the optimal strategy was ignoring emotional faces which might have competed with an increased allocation of attention towards positive stimuli. In fact, there was "only" a specific negativity bias towards fearful faces at baseline, which is a frequent finding in attentional research and might be driven by the evolutionary prioritization of potential threads (Adolphs, 2013; Bertini & Làdavas, 2021; Maratos & Pessoa, 2019; Schindler & Bublatzky, 2020).

Successful control of interfering information in general (Corbetta et al., 2008; Corbetta & Shulman, 2002) as well as the selective allocation of attentional resources in the service of prioritized emotional goals (Cisler & Koster, 2010; Sasse et al., 2014) depend on the ability to control selective attention. Thus, it seems plausible that expectation effects in visual perception rely on participants' ability to exert top-down attentional control. Here we investigated the basic principle of this assumption by assessing individuals' general cognitive ability to exert attentional top-down control over salient distraction with a Singleton visual search task (Theeuwes, 1991, 1992). This task evaluates the specific function of spontaneously focusing on target information in the context of

salient distraction, which is probably involved when people voluntarily control their attention in line with an instructed expectation. Both, behavioral and neural expectation effects were directly correlated with participants' control ability in this task. At the neural level, control ability was associated with a specifically activated network of visuo-spatial orienting and top-down control (Corbetta et al., 2008; Shomstein, 2012), consisting of the right dorsolateral PFC and the bilateral superior parietal lobe (SPL). Together, these findings underline the need to allocate cognitive resources in the service of verbally instructed emotional expectations.

Our findings have strong implications for the understanding of expectation effects in the treatment of mood disorders with antidepressants, both in the context of active placebo treatment and as beneficial "add-on" to effective treatments (Kirsch, 2019; Rief et al., 2016; Rutherford & Roose, 2013). Although prioritizing negative information is a frequent finding in healthy individuals (Bertini & Làdavas, 2021; Maratos & Pessoa, 2019; Norris, 2021; Schindler & Bublatzky, 2020), a pathologically enhanced focus on negative information is a key feature in clinical depression and is crucial for the persistence and prognosis of depressive symptoms (Beck, 1987; LeMoult & Gotlib, 2019; Mennen et al., 2019). It seems therefore intriguing that through verbal instruction alone, attentional preferences for emotional stimuli can be modified together with mood state improvements and positive treatment experience. Whether this also applies to clinical samples and how stable such effects are over time must be addressed in future studies. In this context, the role of cognitive control ability and prefrontal function may be particularly crucial in patients with major depressive disorder who often demonstrate deficits in those domains (Grahek et al., 2019; Millan et al., 2012; Pizzagalli & Roberts, 2022; Snyder, 2013). Our findings suggest that patients with attenuated control abilities may benefit less from verbally instructed expectations as this effect depends on cognitively highly demanding top-down regulation. However, those patients' expectations might be boosted by conditioning, learning procedures and reinstatement of previous positive experiences by which more implicit, cognitively less demanding regulation could be activated (Etkin et al., 2015; Thams & Brassen, 2023). First studies have started to develop such procedures (Peciña et al., 2018; Rutherford et al., 2017; Shafir et al., 2023) and together with an individual assessment of cognitive control ability this may pave the way for an individually tailored intervention in the treatment of affective disorders.

Methods

Participants

In total, fifty-five volunteers aged between 18 and 35 years were invited to participate in the study. All participants met the eligibility criteria for MRI scans, had no history of psychiatric or neurological disorders, and were not currently on any prescribed medication. However, due to technical issues, data from two subjects were excluded, leaving us with fifty-three complete datasets. Three volunteers did not believe they were receiving oxytocin from the beginning, leading to their exclusion from the study. Additionally, one dataset exhibited strong movement artifacts (>50 volumes with scaled variance differences exceeding 2.5 standard deviations as calculated using the TSDiffAna toolbox) and was subsequently removed. This resulted in a final sample size of N = 49. Participants provided informed consent, and ethical approval was obtained from the local ethics committee (Ethikkommission der Ärztekammer Hamburg). All procedures were conducted in adherence to the applicable guidelines and regulations. The study has been registered at the German Clinical Trials Register (https://drks.de/search/en/trial/DRKS00027289).

Experimental protocol and expectation induction

Participants attended the laboratory over a span of three days. The initial day was dedicated to screening including the Singleton task. The subsequent two days involved participants undergoing an emotional interference task within the MRI scanner. Throughout the experiment, mood ratings, expectations, and experiences were recorded (Figure 1).

Expectation effects of intranasal oxytocin were induced utilizing a custom 5-minute video documentary, in which an expert in the field described the foundational mechanisms of the neuropeptide and particularly elaborated on the mood enhancing effects as well as emotional processing, with an emphasis on facial expressions. Additional textual descriptions were provided to ensure a consistent and standardized method of induction. Participants viewed the video on both testing days. Once the video concluded, participants were handed a sealed envelope, which upon opening, disclosed whether they were in the placebo or control condition for that day. The rationale given behind this transparency was to eliminate any potential biases or unexpected placebo effects. Following this, they self-administered a saline nasal spray, dispersing four puffs, two per nostril. On one day, they were informed about the oxytocin (placebo condition) and on the alternate day, they were briefed on the saline nature of the spray (control condition). The sequence of this information delivery was counterbalanced across participants.

Singleton task

On the screening day, participants performed a visual search task on a computer that measures participants' ability to recruit top-down control trial-by-trial to focus on or to ignore salient visual stimuli ("singleton detection mode") (Costello et al., 2010; Sasse et al., 2014; Theeuwes & Burger, 1998). On each trial, nine circles (distractors) and one diamond (target shape) were presented equally spaced around the fixation point. Participants had to indicate as fast and accurately as possible whether the target shape (diamond) contained a horizontal or a vertical line, using dedicated keys on a standard computer keyboard.

There were two conditions indicated by different instructions. In the baseline condition (BC), participants were informed that the target-diamond is always green. Specifically, 50% of all trials contained no particularly salient item (i.e., no singleton, all items green, BC_{all_green}), while in the other 50%, one distractor (circle) was printed in red (singleton distractor, $BC_{singleton_distractor}$). Thus, in the baseline condition participants could completely concentrate on one color. In the Singleton-detection condition (SC), participants were presented with the information, that the color singleton could also be the target. In fact, in ~8% of trials the target-shape was printed in red (singleton target),

in addition to trials with no singleton (SC_{all_green}) and singleton distractors ($SC_{singleton_distractor}$). That is, now the singleton color could not be easily ignored (see Figure 4A). The experiment consisted of 8 blocks (4 for each condition), presented alternately. Each block contained 48 trials, with instructions indicating the condition displayed at the beginning and midway through each block. Each trial began with a fixation period of 1.1 to 1.25 seconds, followed by stimulus presentation for up to 3 seconds, and a random inter-trial interval ranging from 0.2 to 0.8 seconds. Feedback was provided for either a failure to respond within the 3-second window or an incorrect response, with messages displayed as "too slow" or "error", respectively. All trials in which volunteers responded inaccurately or faster than 200 ms were excluded. Trials with a singleton target and the subsequent trial were also excluded due to specific response patterns induced by such trials. The individual singleton score was then calculated using reaction times in the following way:

Singleton score = $[SC_{singleton_distractor} - SC_{all_green}] - [BC_{singleton_distractor} - BC_{all_green}]$

Thus, a higher singleton score implied less flexible attentional control.

Behavioral ratings

Mood ratings were assessed using visual analogue scales (VAS) by which a randomly positioned cursor could be shifted on a scale spanning from 'unhappy' to 'happy' with 200 incremental steps, using the left or right arrow keys on the keyboard. Values were scaled relative to the minimum and maximum provided on each testing day and baseline corrected by subtracting VAS_0 from $VAS_{1\text{-}4}$ values.

Following the nasal spray application, participants indicated their expected positive mood change on an 11-point scale, where 0 represented no expected uplift in mood and 10 signified an expected large mood enhancement. At each study day's conclusion, they conveyed their actual mood shift experience using the same scale (0 for no observed positive mood change, and 10 for a significant positive mood uplift). The rating scales were adapted from (Rief et al., 2017).

Emotional interference task

Participants performed a spatial cueing paradigm modified from a published procedure by Posner et al. (Posner et al., 1980) which has been successfully established in our laboratory (Brassen et al., 2010, 2011). Participants had to respond to left- or right-sided dot-targets, which were preceded by a spatially informative or uninformative cue. After a fixation period of 1000ms, cues were centrally superimposed on neutral, happy, sad, fearful and scrambled faces (5° of visual angle) for a jittered interval of 2050-2500 ms. Then, a dot target appeared on the left or right side (horizontally displaced by 7° of visual angle from the cue position) of the screen for 200 ms. Participants were instructed to respond as quickly and accurately as possible by button press with the left or right index finger according to the side of target presentation. They were told that the faces are irrelevant to the experiment and thoroughly instructed to keep fixation throughout the task. The trial ended with a randomly chosen inter-trial-interval (ITI) of 2000-3500 ms. In two-thirds of all trials, a white arrowhead pointed either to the side where the target appeared (valid, 83%) or to the opposite side (invalid, 17%). The very few invalid trials were implemented to maintain attention and were not included in further data analyses. In the other one-third of all trials, a neutral, non-spatial cue indicated that the target will be equally likely to occur at both possible locations.

A total of 360 trials were administered across three sessions, with 1-minute break intervals in between. Participants initiated the first session and concluded each subsequent one by completing the VAS mood assessment. The design was fully counterbalanced with pseudo-randomized trial order. Altogether, there were 18 facial identities (9 female) chosen from the well-validated Karolinska Directed Emotional Faces set (Lundqvist et al., 1998). Each identity displayed each of the

four emotional expressions, and scrambled pictures were derived by randomly shuffling the pixels of the neutral face. Each picture was displayed 4 times in total. Prior to scanning, participants familiarized themselves with the task in a training session outside the scanner. After scanning, participants rated the valence and arousal (range: 1-4) of all faces that were presented during the task to control for potential individual differences in valence categorization and to characterize potential differences in arousal ratings between facial expressions and sessions. Both the cognitive control and the emotional interference task, along with the arousal assessments, were executed using the Psychophysics Toolbox (PTB, 3.0.14) for MATLAB.

Eye movement recording

We conducted an online eye position tracking during fMRI scanning with an infrared eye-tracker (Eyelink 1000, SR Research). The camera was focused on the right eye. Participants completed a calibration procedure before each session, fixating on 9 calibration points. The viewing distance was 163 cm in total. Gaze positions were sampled at 1000 Hz, allowing accurate saccade and fixation tracking. Stimulus onsets were recorded in parallel to the eye-tracking data.

Eye-tracking data were first parsed into rapid eye movements exceeding a velocity of 30°/s or an acceleration of 8000°/s^2 (saccades) and relatively stable gaze periods between the saccades (fixations). Subsequently, we subtracted the last 300 ms before stimulus onset from fixation coordinates to relate individual gaze positions to baseline values where the centrally presented cross should be fixated. To ensure that participants adhered to this instruction, we used an iterative outlier detection algorithm on x- and y-coordinates of baseline values (End & Gamer, 2017). Specifically, the highest and lowest values of baseline data were temporally removed from the data set, and it was tested whether they deviated more than three SDs from the mean of the remaining distribution. If one or both values met this criterion, they were marked as outliers and excluded from the data set, if not, they were returned. This was iteratively repeated until no further x- and y-values had to be removed. Trials with blinks during the baseline period were also removed.

After trial selection and drift correction, we identified all fixations that fell on the left ($x < 0$) and right side of the screen ($x > 0$) and calculated the total dwell time on both sides during the entire cueing period before the target appeared. From these values, we derived the proportion of dwell time on the screen side that was cued (for directional cues) or on the right side of the screen (for neutral cues) by dividing dwell times by the total duration of all valid fixations in this period. These values were averaged for all trials of the same condition and compared using a 2 (day: placebo vs. control) × 5 (face: fearful, happy, sad, neutral, scrambled) × 2 (cue: non-directed vs. directed) ANOVA. For illustrative purposes, we also calculated trajectories of horizontal gaze displacements from fixation coordinates in bins of 10 ms for the whole cueing period.

Due to significant data loss, we could only analyze the eye-tracking data of 27 participants who had at least 5 trials with valid eye-tracking data in each cell of the experimental paradigm. On average, 259.61 (SD = 72.27) trials were available for each participant and day.

MR image acquisition

Imaging data was acquired on a Siemens PRISMA 3T scanner (Erlangen, Germany) using a 64-channel head coil. Multiband echo-planar imaging sequences were applied to obtain the functional imaging data. Each volume of the experimental data contained 66 slices (voxel size of 1.5 × 1.5 × 1.5 mm) and was angled to the anterior to posterior commissure (AC-PC) line (repetition time (TR) = 1.55 s, echo time (TE) = 29 ms, flip angle = 70°, field of view (FoV) = 225 mm, multiband mode with 3 bands). A T_1-weighted structural image was also acquired for each subject with a magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence with 240 slices and a voxel size of 1 × 1 × 1 mm for functional preprocessing.

Image preprocessing and analysis

Preprocessing and image analysis were performed utilizing SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK) and custom scripts in MATLAB. Each functional volume was corrected for rigid body motion as well as susceptibility artefacts using realign and unwarp. The individual T1 structural image was coregistered to the average functional image produced during the realignment procedure. The functional images were subsequently normalized and smoothed with a 4-mm full-width at half maximum Gaussian kernel. The TsDiffAna toolbox (http://www.fil.ion.ucl.ac.uk/spm/ext/#TSDiffAna) was utilized for image diagnostics to control for artefacts.

Statistical analysis was conducted using a two-level random effects approach employing the general linear model as implemented in SPM12. At the single subject level, onsets of all 10 conditions (5 emotions × 2 cues) were modeled as separate regressors convolving delta functions with a canonical hemodynamic response function. Additionally, a regressor of no interest representing invalid and incorrect responses was defined. Data from the placebo and control condition were defined as separate sessions and entered into a single model. Furthermore, to enhance signal-to-noise ratio in our dataset, we utilized the GLMdenoise toolbox (https://kendrickkay.net/GLMdenoise/, Version 1.4). Noise regressors were obtained through principal component and cross-validation analyses on voxel time-series that were determined by a preliminary first-level-model fit to be unrelated to the experimental paradigm. Individual noise regressors were entered as regressors of no interest into the first-level model.

For each subject, contrast images for regressors of interest were then entered into second-level random-effect ANOVA models including the factors stimulus (all emotions for confirmatory analyses, happy and fearful faces for target analyses), cue and session. For correlation analysis including the singleton score, a simple regression model was used.

To assess the functional connectivity between the identified PFC cluster with other brain regions, we conducted a generalized psychophysiological interaction analysis (gPPI toolbox; http://www.nitrc.org/projects/gppi). Accordingly, we extracted the time course from 2mm spheres around individual peaks within the identified cluster. On the first level, adjusted time series were modeled as regressors together with vectors of condition-onsets and the interaction of both. Specific interaction contrast images (i.e., fearful/happy, non-directed/directed, placebo/control) were then entered into a second level factorial design.

We report results corrected for FWE due to multiple comparisons. We conducted this correction at the peak level within small volume ROIs for which we had an a priori hypothesis or when passing a whole-brain corrected cluster threshold of FWE $< .05$ (cluster forming threshold $p < .001$ uncorrected).

Based on our main interest in placebo-related top-down regulation during emotional processing, we focused on prefrontal areas and the amygdala. Specifically, we created ROIs based on functional clusters ($p < 0.05$ FDR corrected) extracted from meta-analyses conducted on the neurosynth.org platform, using the search terms “DLPFC” (489 studies) and “VMPFC” (199 studies; August 2023). Due to strong overlaps of placebo results in the vmPFC and ACC, the vmPFC was extended by the anatomical mask of the anterior cingulate cortex provided by the AAL3 atlas (Rolls 2020). The bilateral anatomical mask of the amygdala (AAL3) was defined as our third ROI.

References

Abrams, K., Kushner, M., Lisdahl Medina, K., & Voight, A. (2001). The pharmacologic and expectancy effects of alcohol on social anxiety in individuals with social phobia. *Drug and Alcohol Dependence*, *64*(2), 219–231. https://doi.org/10.1016/S0376-8716(01)00125-9

Adolphs, R. (2013). The Biology of Fear. *Current Biology*, *23*(2), R79–R93. https://doi.org/10.1016/j.cub.2012.11.055

Anna, J. A., Julia, R., Julia, W., Lea, R., & Winfried, R. (2019). Placebo mechanisms in depression: An experimental investigation of the impact of expectations on sadness in female participants. *Journal of Affective Disorders*, *256*, 658–667. https://doi.org/10.1016/j.jad.2019.06.070

Atlas, L. Y., & Wager, T. D. (2014). A Meta-analysis of Brain Mechanisms of Placebo Analgesia: Consistent Findings and Unanswered Questions. In F. Benedetti, P. Enck, E. Frisaldi, & M. Schedlowski (Eds.), *Placebo* (pp. 37–69). Springer. https://doi.org/10.1007/978-3-662-44519-8_3

Baker, J., Gamer, M., Rauh, J., & Brassen, S. (2022). Placebo induced expectations of mood enhancement generate a positivity effect in emotional processing. *Scientific Reports*, *12*(1), 5345. https://doi.org/10.1038/s41598-022-09342-2

Balodis, I. M., Wynne-Edwards, K. E., & Olmstead, M. C. (2011). The stress–response-dampening effects of placebo. *Hormones and Behavior*, *59*(4), 465–472. https://doi.org/10.1016/j.yhbeh.2011.01.004

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy*, *1*, 5–37.

Benedetti, F. (2014). Placebo effects: From the neurobiological paradigm to translational implications. *Neuron*, *84*(3), 623–637. https://doi.org/10.1016/j.neuron.2014.10.023

Bertini, C., & Làdavas, E. (2021). Fear-related signals are prioritised in visual, somatosensory and spatial systems. *Neuropsychologia*, *150*, 107698. https://doi.org/10.1016/j.neuropsychologia.2020.107698

Brassen, S., Gamer, M., & Büchel, C. (2011). Anterior Cingulate Activation Is Related to a Positivity Bias and Emotional Stability in Successful Aging. *Biological Psychiatry*, *70*(2), 131–137. https://doi.org/10.1016/j.biopsych.2010.10.013

Brassen, S., Gamer, M., Rose, M., & Büchel, C. (2010). The influence of directed covert attention on emotional face processing. *NeuroImage*, *50*(2), 545–551. https://doi.org/10.1016/j.neuroimage.2009.12.073

Buhle, J. T., Stevens, B. L., Friedman, J. J., & Wager, T. D. (2012). Distraction and Placebo: Two Separate Routes to Pain Control. *Psychological Science*, *23*(3), Article 3. https://doi.org/10.1177/0956797611427919

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), Article 2. https://doi.org/10.1016/j.cpr.2009.11.003

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), Article 3. https://doi.org/10.1016/j.neuron.2008.04.017

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Costello, M. C., Madden, D. J., Shepler, A. M., Mitroff, S. R., & Leber, A. B. (2010). Age-Related Preservation of Top-Down Control Over Distraction in Visual Search. *Experimental Aging Research*, *36*(3), 249–272. https://doi.org/10.1080/0361073X.2010.484719

Crichton, F., Dodd, G., Schmid, G., Gamble, G., Cundy, T., & Petrie, K. J. (2014). The power of positive and negative expectations to influence reported symptoms and mood during exposure to wind farm sound. *Health Psychology*, *33*, 1588–1592. https://doi.org/10.1037/hea0000037

Ellingsen, D.-M., Wessberg, J., Eikemo, M., Liljencrantz, J., Endestad, T., Olausson, H., & Leknes, S. (2013). Placebo improves pleasure and pain through opposite modulation of sensory processing. *Proceedings of the National Academy of Sciences*, *110*(44), Article 44. https://doi.org/10.1073/pnas.1305050110

End, A., & Gamer, M. (2017). Preferential Processing of Social Features and Their Interplay with Physical Saliency in Complex Naturalistic Scenes. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00418

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, *16*(11), 693–700. https://doi.org/10.1038/nrn4044

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), Article 1. https://doi.org/10.1038/s41386-021-01132-0

Geers, A. L., Faasse, K., Guevarra, D. A., Clemens, K. S., Helfer, S. G., & Colagiuri, B. (2021). Affect and emotions in placebo and nocebo effects: What do we know so far? *Social and Personality Psychology Compass*, *15*(1), e12575. https://doi.org/10.1111/spc3.12575

Geuter, S., Koban, L., & Wager, T. D. (2017). The Cognitive Neuroscience of Placebo Effects: Concepts, Predictions, and Physiology. *Annual Review of Neuroscience*, *40*, 167–188. https://doi.org/10.1146/annurev-neuro-072116-031132

Grahek, I., Shenhav, A., Musslick, S., Krebs, R. M., & Koster, E. H. W. (2019). Motivation and Cognitive Control in Depression. *Neuroscience and Biobehavioral Reviews*, *102*, 371–381. https://doi.org/10.1016/j.neubiorev.2019.04.011

Haas, J. W., Rief, W., Glombiewski, J. A., Winkler, A., & Doering, B. K. (2020). Expectation-induced placebo effect on acute sadness in women with major depression: An experimental investigation. *Journal of Affective Disorders*, *274*, 920–928. https://doi.org/10.1016/j.jad.2020.05.056

Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 4–26. https://doi.org/10.1038/npp.2009.129

Hahn, A., Göhler, A. C., Hermann, C., & Winkler, A. (2022). Even when you know it is a placebo, you experience less sadness: First evidence from an experimental open-label placebo investigation. *Journal of Affective Disorders*, *304*, 159–166. https://doi.org/10.1016/j.jad.2022.02.043

Kirsch, I. (2019). Placebo Effect in the Treatment of Depression and Anxiety. *Frontiers in Psychiatry*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00407

Koban, L., Jepma, M., Geuter, S., & Wager, T. D. (2017). What's in a word? How instructions, suggestions, and social information change pain and emotion. *Neuroscience and Biobehavioral Reviews*, *81*(Pt A), 29–42. https://doi.org/10.1016/j.neubiorev.2017.02.014

Lapate, R. C., Ballard, I. C., Heckner, M. K., & D'Esposito, M. (2022). Emotional Context Sculpts Action Goal Representations in the Lateral Frontal Pole. *Journal of Neuroscience*, *42*(8), 1529–1541. https://doi.org/10.1523/JNEUROSCI.1522-21.2021

LeMoult, J., & Gotlib, I. H. (2019). Depression: A cognitive perspective. *Clinical Psychology Review*, *69*, 51–66. https://doi.org/10.1016/j.cpr.2018.06.008

Leuchter, A. F., Hunter, A. M., Tartter, M., & Cook, I. A. (2014). Role of pill-taking, expectation and therapeutic alliance in the placebo response in clinical trials for major depression. *The British Journal of Psychiatry*, *205*(6), 443–449. https://doi.org/10.1192/bjp.bp.113.140343

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces—KDEF, CD ROM.* Department of Clinical Neuroscience, Psychology section.

Maratos, F. A., & Pessoa, L. (2019). Chapter 6 - What drives prioritized visual processing? A motivational relevance account. In N. Srinivasan (Ed.), *Progress in Brain Research* (Vol. 247, pp. 111–148). Elsevier. https://doi.org/10.1016/bs.pbr.2019.03.028

Mennen, A. C., Norman, K. A., & Turk-Browne, N. B. (2019). Attentional bias in depression, Understanding mechanisms to improve training and treatment. *Current Opinion in Psychology*, *29*, 266–273. https://doi.org/10.1016/j.copsyc.2019.07.036

Meyer, B., Yuen, K. S. L., Ertl, M., Polomac, N., Mulert, C., Büchel, C., & Kalisch, R. (2015). Neural mechanisms of placebo anxiolysis. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(19), 7365–7373. https://doi.org/10.1523/JNEUROSCI.4793-14.2015

Meyer, B., Yuen, K. S. L., Saase, V., & Kalisch, R. (2019). The Functional Role of Large-scale Brain Network Coordination in Placebo-induced Anxiolysis. *Cerebral Cortex (New York, N.Y.: 1991)*, *29*(8), 3201–3210. https://doi.org/10.1093/cercor/bhy188

Millan, M. J., Agid, Y., Brüne, M., Bullmore, E. T., Carter, C. S., Clayton, N. S., Connor, R., Davis, S., Deakin, B., DeRubeis, R. J., Dubois, B., Geyer, M. A., Goodwin, G. M., Gorwood, P., Jay, T. M., Joëls, M., Mansuy, I. M., Meyer-Lindenberg, A., Murphy, D., … Young, L. J. (2012). Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. *Nature Reviews. Drug Discovery*, *11*(2), 141–168. https://doi.org/10.1038/nrd3628

Norris, C. J. (2021). The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. *Social Neuroscience*, *16*(1), 68–82. https://doi.org/10.1080/17470919.2019.1696225

Peciña, M., Heffernan, J., Wilson, J., Zubieta, J. K., & Dombrovski, A. Y. (2018). Prefrontal expectancy and reinforcement-driven antidepressant placebo effects. *Translational Psychiatry*, *8*(1), Article 1. https://doi.org/10.1038/s41398-018-0263-y

Petrovic, P., Dietrich, T., Fransson, P., Andersson, J., Carlsson, K., & Ingvar, M. (2005). Placebo in Emotional Processing—Induced Expectations of Anxiety Relief Activate a Generalized Modulatory Network. *Neuron*, *46*(6), 957–969. https://doi.org/10.1016/j.neuron.2005.05.023

Pizzagalli, D. A., & Roberts, A. C. (2022). Prefrontal cortex and depression. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *47*(1), 225–246. https://doi.org/10.1038/s41386-021-01101-7

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, *109*(2), 160–174.

Rief, W., Barsky, A. J., Bingel, U., Doering, B. K., Schwarting, R., Wöhr, M., & Schweiger, U. (2016). Rethinking psychopharmacotherapy: The role of treatment context and brain plasticity in antidepressant and antipsychotic interventions. *Neuroscience and Biobehavioral Reviews*, *60*, 51–64. https://doi.org/10.1016/j.neubiorev.2015.11.008

Rief, W., Nestoriuc, Y., Weiss, S., Welzel, E., Barsky, A. J., & Hofmann, S. G. (2009). Meta-analysis of the placebo response in antidepressant trials. *Journal of Affective Disorders*, *118*(1), 1–8. https://doi.org/10.1016/j.jad.2009.01.029

Rief, W., Shedden-Mora, M. C., Laferton, J. A. C., Auer, C., Petrie, K. J., Salzmann, S., Schedlowski, M., & Moosdorf, R. (2017). Preoperative optimization of patient expectations improves long-term outcome in heart surgery patients: Results of the randomized controlled PSY-HEART trial. *BMC Medicine*, *15*(1), 4. https://doi.org/10.1186/s12916-016-0767-3

Rutherford, B. R., & Roose, S. P. (2013). A Model of Placebo Response in Antidepressant Clinical Trials. *The American Journal of Psychiatry*, *170*(7), 723–733. https://doi.org/10.1176/appi.ajp.2012.12040474

Rutherford, B. R., Wall, M. M., Brown, P. J., Choo, T.-H., Wager, T. D., Peterson, B. S., Chung, S., Kirsch, I., & Roose, S. P. (2017). Patient Expectancy as a Mediator of Placebo Effects in Antidepressant Clinical Trials. *American Journal of Psychiatry*, *174*(2), 135–142. https://doi.org/10.1176/appi.ajp.2016.16020225

Sasse, L. K., Gamer, M., Büchel, C., & Brassen, S. (2014). Selective Control of Attention Supports the Positivity Effect in Aging. *PLOS ONE*, *9*(8), e104180. https://doi.org/10.1371/journal.pone.0104180

Schaefer, M., Kühnel, A., Schweitzer, F., Enge, S., & Gärtner, M. (2023). Neural underpinnings of open-label placebo effects in emotional distress. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *48*(3), 560–566. https://doi.org/10.1038/s41386-022-01501-3

Schienle, A., Unger, I., & Schwab, D. (2022). Changes in neural processing and evaluation of negative facial expressions after administration of an open-label placebo. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-10567-4

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *130*, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Schwarz, K. A., Pfister, R., & Büchel, C. (2016). Rethinking Explicit Expectations: Connecting Placebos, Social Cognition, and Contextual Perception. *Trends in Cognitive Sciences*, *20*(6), 469–480. https://doi.org/10.1016/j.tics.2016.04.001

Sezer, D., Locher, C., & Gaab, J. (2022). Deceptive and open-label placebo effects in experimentally induced guilt: A randomized controlled trial in healthy subjects. *Scientific Reports*, *12*(1), 21219. https://doi.org/10.1038/s41598-022-25446-1

Shafir, R., Israel, M., & Colloca, L. (2023). Harnessing the placebo effect to enhance emotion regulation effectiveness and choice. *Scientific Reports*, *13*(1), 2373. https://doi.org/10.1038/s41598-023-29045-6

Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: Top-down and bottom-up attentional control. *Frontiers in Integrative Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnint.2012.00038

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. https://doi.org/10.1037/a0028727

Song, S., Zilverstand, A., Song, H., d'Oleire Uquillas, F., Wang, Y., Xie, C., Cheng, L., & Zou, Z. (2017). The influence of emotional interference on cognitive control: A meta-analysis of neuroimaging studies using the emotional Stroop task. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-02266-2

Stone, M. B., Yaseen, Z. S., Miller, B. J., Richardville, K., Kalaria, S. N., & Kirsch, I. (2022). Response to acute monotherapy for major depressive disorder in randomized, placebo controlled trials submitted to the US Food and Drug Administration: Individual participant data analysis. *BMJ*, *378*, e067606. https://doi.org/10.1136/bmj-2021-067606

Thams, F., & Brassen, S. (2023). The need to change: Is there a critical role of midlife adaptation in mental health later in life? *ELife*, *12*, e82390. https://doi.org/10.7554/eLife.82390

Thase, M. E., Larsen, K. G., & Kennedy, S. H. (2011). Assessing the 'true' effect of active antidepressant therapy v. placebo in major depressive disorder: Use of a mixture model. *The British Journal of Psychiatry*, *199*(6), 501–507. https://doi.org/10.1192/bjp.bp.111.093336

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/bf03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/bf03211656

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1342–1353. https://doi.org/10.1037/0096-1523.24.5.1342

van der Meulen, M., Kamping, S., & Anton, F. (2017). The role of cognitive reappraisal in placebo analgesia: An fMRI study. *Social Cognitive and Affective Neuroscience*, *12*(7), 1128–1137. https://doi.org/10.1093/scan/nsx033

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., Kosslyn, S. M., Rose, R. M., & Cohen, J. D. (2004). Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain. *Science*, *303*(5661), 1162–1167. https://doi.org/10.1126/science.1093065

Zhang, W., Qin, S., Guo, J., & Luo, J. (2011). A follow-up fMRI study of a transferable placebo anxiolytic effect. *Psychophysiology*, *48*(8), 1119–1128. https://doi.org/10.1111/j.1469-8986.2011.01178.x

Zunhammer, M., Spisák, T., Wager, T. D., & Bingel, U. (2021). Meta-analysis of neural systems underlying placebo analgesia from individual participant fMRI data. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21179-3

End notes

Author contributions

Concept and design: S.B., Acquisition of data: J.R., A.M., Analysis or interpretation of data: J.R., M.G., C.B., W.R., S.B., Drafting of the manuscript: S.B., J.R., A.M., Critical revision of the manuscript: All authors.

Acknowledgements

We thank F. Thams, J. Baker and E. Shin for their assistance with data collection. The work was funded by grant CRC 289 Treatment Expectation, Project-ID 422744262, from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)

Competing financial interests: The authors declare no competing financial interests.