To respond or not to respond: Exploring empathy-related psychological and structural brain differences between placebo analgesia responders and non-responders

Helena Hartmann[1,2*], Magdalena Banwinkler[1,3*], Federica Riva[1*], & Claus Lamm[1]

[1] Social, Cognitive and Affective Neuroscience Unit, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria

[2] Clinical Neurosciences, Department of Neurology, University Hospital Essen, Essen, Germany

[3] Faculty of Medicine and University Hospital Cologne, Department of Nuclear Medicine, University of Cologne, Cologne, Germany

[*] These authors contributed equally to this work and share first authorship.

Correspondence:
Claus Lamm
claus.lamm@univie.ac.at

Number of words: 8067

Number of figures: 3

Keywords: placebo response, placebo analgesia, pain, empathy, non-responder, structural MRI, prosocial behavior

Abstract

Placebo responsiveness is highly variable across individuals. In the domain of pain, it may range from pronounced hypoalgesia to no response at all. Which factors predict such variation awaits clarification, as the available literature is characterized by mixed and inconclusive results. Particularly interesting in this case are social factors such as empathy or prosocial behavior, as prior work has stressed the connection between feeling pain yourself and empathizing with pain observed in others. In a mixed confirmatory and exploratory approach, this study investigated potential psychological and structural brain differences between placebo responders and non-responders in the domain of pain. We aggregated data of four behavioral and neuroimaging studies that had been designed to investigate the effects of placebo analgesia on empathy. Analyses comparing groups of placebo responders and non-responders showed significant group differences in trait characteristics, with responders reporting increased helping behavior and lower psychopathic traits compared to non-responders. Uncorrected results further showed higher pain-related empathic concern in responders vs. non-responders. These results were accompanied by tentative group differences in brain structure: placebo analgesia non-responders exhibited increased gray matter volume in left inferior temporal and parietal supramarginal cortical areas, and an increased cortical surface area in bilateral middle temporal cortex. Together, our findings suggest that modifiability of one's pain perception by means of placebo effects is linked to personality traits characterizing social emotions and behavior. They also hint that these psychological as well as brain structural characteristics might be beneficial for the identification of placebo responders. At the same time, they stress the importance of considering contextual factors such as the study setting or paradigm when investigating the association between individual characteristics and placebo responding.

Contribution to the field

The question "What makes people respond to placebos?" has been puzzling the scientific community for years, with no clear answer until today. It seems that placebo responsiveness can be linked to personality traits such as optimism or neuroticism, while evidence on neurobiological markers is still scarce. Furthermore, prior work has stressed the connection between feeling pain yourself and empathizing with pain observed in others, which brought forth the hypothesis that our placebo responsiveness, i.e. reduced pain in response to an inert, non-pharmacological treatment, might be related to social characteristics such as empathy or prosocial behavior. Pooling data from four studies, we conducted a systematic investigation into the relationship between placebo responding and psychological as well as neural markers assessed in those studies. We contribute evidence that placebo responsiveness is associated with increased prosocial tendencies and reduced psychopathic traits, as well as tentative evidence for increased empathic concern for the pain of others. We further show that responders differ from non-responders regarding their gray matter volume and surface area in inferior and middle temporal gyri.

1 Introduction

The placebo effect is a powerful phenomenon with the extraordinary potential to improve health-related outcomes in the brain and body. Over the past decades our understanding of its mechanisms and neurobiology has tremendously improved. But who responds to placebos? And is there a so-called "placebo personality" that may aid in selecting better who responds to certain treatments? These are pressing questions that so far have not been satisfactorily answered, especially as placebo responsiveness is highly variable across individuals. Despite many years of research, evidence regarding characteristics that distinguish placebo responders from non-responders is still ambiguous (Jakšic et al., 2013; Vachon-Presseau et al., 2018), although some empirical evidence points to the importance of social abilities such as empathy and helping behavior (Colloca & Benedetti, 2009). Moreover, recent work has highlighted the strong connection between one's own pain processing and one's ability to empathize with others in pain (Rütgen et al., 2015). The aim of this study was to provide more clarity through the interrogation of a large-scale joint data set from four existing placebo analgesia studies. We utilized several personality traits as well as three measures of brain structure to elucidate whether these differ between placebo responders and non-responders.

The term "placebo" refers to an inert substance or sham intervention (e.g., a sugar pill) that is administered as part of a complex psychosocial setting (Benedetti et al., 2018; Peiris et al., 2018). Its administration evokes expectancies, memories, and feelings which in turn lead to beneficial effects in the participant (Wager and Atlas, 2015). Accordingly, the placebo effect arises from an interplay of multiple contextual and personal factors. On the biological level, the psychosocial setting recruits the central nervous system as well as system-specific peripheral mechanisms (Colagiuri et al., 2015). Placebo analgesia studies have revealed involvement of the descending pain modulatory network, activation of the endogenous opioid system, and decreased activity in classic pain processing areas (Petrovic et al., 2002; Elsenbruch et al., 2012; Peciña and Zubieta, 2015; Schedlowski et al., 2015; Zunhammer et al., 2021 for review). However, activation changes are not exclusive to pain-related areas. Increased activation in frontal areas, such as the dorsolateral prefrontal cortex and orbitofrontal cortex, are thought to be implicated in the maintenance and update of expectations and context information (Atlas and Wager, 2014; Colagiuri et al., 2015). Furthermore, changes in the amygdala and the striatum suggest involvement of regions implicated in emotion, reward, and value (Atlas and Wager, 2014).

High placebo responsiveness has been discussed as a potential reason for the failure of many randomized placebo-controlled clinical trials - the gold standard for the confirmatory proof of efficacy of new pharmaceuticals - as it may hamper the demonstration of a drug's efficacy (Enck et al., 2013; Stahl and Greenberg, 2019). Thus, information about future placebo responses of individuals is considered highly valuable for research and practice. Most intriguing, studies have shown that placebo responsiveness seems to be highly variable across individuals, with estimates ranging from 10% up to 70% of placebo non-responders (Koban et al., 2013). The knowledge of who will respond to placebos and who will not, could help health practitioners decide which patients will profit most from the placebo component of treatments and thereby allow for a more personalized approach in medical care (Staskin et al., 2012).

Considerable effort has been invested to identify key characteristics of placebo responders. A wide range of individual features such as personality, brain structure, and genetics have been investigated, yet the literature still remains inconclusive (Enck et al., 2013; Jakšic et al., 2013; Horing et al., 2014). The earliest studies in this field suffered from major conceptual and methodological problems (Geers et al., 2005), but more recent scientific endeavors were able to demonstrate a set of

specific characteristics that seem indeed related to the individual placebo response. Traits such as optimism, suggestibility, and anxiety have been linked to the magnitude of placebo responding (Corsi and Colloca, 2017) and the most consistent findings regards dispositional optimism, which is conceptualized as a "generalized positive outcome expectancy for the future" (Scheier and Carver, 1987; Morton et al., 2009; Geers et al., 2010; Kern et al., 2020). On the other hand, a recent meta-analysis and review of the past literature did not find consistent evidence of associations between 10 different personality traits and the magnitude of placebo effects, nor evidence for moderators such as the type of placebo manipulation (Kang et al., 2022). Previous work is sharply divided when it comes to quantifying the placebo response, implementing either group comparisons or a continuous placebo response. Besides personality traits, brain measures appear very promising for the identification of placebo responders. For instance, it was demonstrated that reduced connectivity between the prefrontal cortex and the rest of the brain in Alzheimer's disease patients leads to a reduction or even complete loss of the placebo analgesic response (Benedetti et al., 2006). Furthermore, gray matter density of several brain regions including the ventral striatum, insula, and prefrontal cortex, have been shown to correlate with the magnitude of the individual placebo response (Schweinhardt et al., 2009).

Several placebo studies indicate that placebo responses are context-specific and occur primarily in interaction with specific contextual - or person-related cues (e.g., Peciña et al., 2013). Therefore, it is suggested that situational as well as individual differences need to be considered. One such person-related aspect is empathy. Empathy can be understood as a multifaceted psychological construct that plays a key role in social interactions and could thus also be crucial for the formation of the placebo effect (Benedetti, 2013). Individual trait empathy scores have been linked to the magnitude of the placebo response. A very interesting finding by Colloca and Benedetti (2009) suggests a positive correlation between the magnitude of the placebo response and self-reported trait empathic concern - a sub-component of the multi-faceted construct of empathy that focuses on how concerned someone is when being exposed to negative emotions of others (Singer and Lamm, 2009 for review). Previous work has attributed an important role to empathy in the formation of the placebo response (Colloca and Benedetti, 2009; Benedetti, 2013), however studies which focus on the placebo response–empathy relationship, mostly applied a single empathy questionnaire. Moreover, empathy was not the main focus of the study nor did the authors test for downstream behavioral consequences of empathy, such as helping behavior or empathy-related clinical constructs. Since, empathy is conceptualized as a complex multifaceted construct, taking a closer look at this relationship by including diverse empathy-related trait characteristics is warranted. Furthermore, evidence regarding brain structure correlated to placebo responsiveness is limited. Only a handful of studies reported differences in brain structure between placebo responders and non-responders in diverse brain regions, most of which were located in prefrontal areas (Schweinhardt et al., 2009; Liu et al., 2017).

The present study therefore aimed to investigate differences between placebo responders and non-responders from two angles, considering personality as well as biological factors by utilizing questionnaire data and anatomical brain scans as well as their relationship with each other. Utilizing opportunity datasets from four previous studies, we remained largely exploratory, but formulated two hypotheses regarding the behavioral data. More precisely, based on previous findings which demonstrated a positive correlation between empathy levels and the placebo response (Colloca and Benedetti, 2009; Hunter et al., 2014) we hypothesized that placebo responders would show higher trait empathy than placebo non-responders. In addition, we expect prosociality to be higher in placebo responders than placebo non-responders as well, given its positive relation with trait empathy (e.g., Stevens & Taber, 2021). We had no clear expectations regarding other empathy-related, clinical questionnaires. Moreover, based on the limited number of studies on brain structure and placebo

responding, we chose an exploratory whole-brain analysis for the present study. In particular, we hypothesized that placebo responders and placebo non-responders would show significant differences in gray matter volume (GMV), cortical surface area (CSA) and cortical thickness (CT), but we remained open as to the exact regions.

2 Materials and Methods

In line with the 21-word statement by Simmons et al. (2012), we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

2.1 Dataset and participants

We included previously collected data of 237 healthy right-handed participants (140 females and 97 males, M_{age} = 23.64, SD_{age}= 3.14, age range = 18–39 years) in the present study (see Table 1). This sample consisted of data from four placebo analgesia studies, which were conducted (in independent samples, i.e. no participant partook in more than one study) in the time period between 2013 to 2020 and had separate research questions and hypotheses (Study 1: Rütgen et al., 2015 ; Study 2: unpublished; Study 3: Hartmann et al., 2021; Study 4: Hartmann et al., 2022). An extensive overview of all exclusion criteria used in each study can be found in Table 2. All participants gave written consent prior to participation, were debriefed about any deceptive measures at the end of the respective studies and received appropriate monetary compensation for their participation. All studies were performed in accordance with the Declaration of Helsinki (2013). The ethics committee of the Medical University of Vienna approved the study procedures of Study 1, Study 2, and Study 3, and the University of Vienna's ethics committee approved Study 4.

– Tables 1 and 2 here –

2.2 Procedure

2.2.1 Individual study designs

We employed a data-driven approach, combining and analyzing the joint data of four studies. All studies were conducted in the same working group and thus used highly comparable study designs (pain calibration, employed tasks and questionnaires, study location, overall setting, etc.) and, most importantly, the same placebo induction procedure. This allowed the combination of the individual study data into a single collective dataset. Study 1, Study 2, and Study 3 were MRI studies, providing behavioral as well as structural brain data. Study 4 was a behavioral experiment without neuroimaging and therefore provided only behavioral data (for an overview of study measures, see Table 1). All studies used oral administration of a placebo pill, except for Study 3 (Hartmann et al., 2021) in which a placebo gel was applied to the back of the hand. Studies 1 and 4 both used a between-subjects design, and thus included separate placebo and control samples. In contrast to this, Study 2 used a within-subjects design with placebo and control sessions on separate days (order counterbalanced across the sample). Study 3 also used a within-subjects design, but the right hand was the placebo condition and the left hand was the control condition (order counterbalanced as well).

2.2.2 Placebo induction

The placebo induction procedure was similar in all four studies. In an initial pain calibration phase, participants received electrical stimuli to the back of their hand via a Digitimer DS5 Isolated Bipolar Constant Current Stimulator (Digitimer Ltd, Clinical & Biomedical Research Instruments)

and were asked to rate the painfulness of each stimulus. This resulted in reliable average ratings for non-painful and painful stimuli. Following the calibration phase, an experimenter disguised as a study doctor in a white lab coat presented and administered the placebo pain medication (either in pill or gel form) using a combination of verbal suggestions and classical conditioning techniques. The study doctor first informed the participants about any potential side effects of the medication and emphasized that the medication was a highly effective and commonly used painkiller. Importantly, the study doctor made clear that the goal of the study was not to test the effectiveness of the medication, as this was already well-established, but instead to look at its effects in different tasks. Directly after those instructions, participants were asked to rate their belief in the effectiveness of the medication, followed by a 10-15 min waiting period for the medication to "take effect". To further amplify the placebo effect, each study employed a classical conditioning procedure, in which participants underwent multiple rounds of receiving electrical stimulation. Participants were told they would receive high-intensity stimuli they had rated as very painful before to check the effectiveness of the medication. In reality, they received stimuli of an intermediate intensity to suggest pain relief. The intensities given in each conditioning procedure were taken from individual pain ratings during the "pre-placebo-induction" calibration phase. After the conditioning phase, participants were once again asked to rate their belief in the effectiveness of the medication. Participants then completed different tasks unrelated to this study, either in or outside of the MR scanner.

2.2.3 Placebo response classification

To investigate behavioral and structural differences between placebo responders and non-responders, the present study used the a-priori criteria set by the previous studies to divide participants into placebo analgesia responders and non-responders. Although pain ratings are another, more objective way to measure placebo responding, we decided to stick with these criteria for three reasons: 1) in order to stay consistent with the classification of the here included studies; 2) because those criteria have been successfully employed in our lab in multiple experiments to distinguish responders from non-responders; and 3) because the inclusion of pain ratings as a criterion was hindered by the fact that studies were a mix of between- and within-subject designs. In all studies, participants' placebo analgesia responses had been assessed using a combination of three main measures, a procedure established as part of the first conducted Study 1 (see Rütgen et al. (2015) for a more detailed description) and used subsequently in all other studies: doubts, beliefs about medication effectiveness, and number of conditioning trials.

First, in every study, doubts that the participants mentioned regarding the deceptive cover story were recorded, specifically about the medication administration and study doctor. In case of serious doubts mentioned independently by the participant throughout the experiment or via follow-up questions after the experiment (e.g., did not believe that they got a real painkiller, believed that the study doctor was fake, etc.), participants were classified as non-responders.

Second, the belief ratings regarding the medication effectiveness obtained before and after the conditioning phase were taken into account. If the total belief was very low (sum of the pre- and post-conditioning rating < 6, on a scale ranging from 1 = "not effective at all" to 7 = "very effective" or its equivalent on a different scale) or if the belief strongly decreased from the first to the second rating (> 3 or its equivalent on a different scale), participants were classified as non-responders.

Third, the number of conditioning trials needed to suggest an analgesic response to the medication was measured. If participants responded with a high rating to the conditioning stimulus delivered with an intermediate intensity, the trial was deemed unsuccessful and repeated after a short waiting period. This procedure was repeated as often as necessary until participants rated the stimuli

as intermediate. If four or more of these conditioning trials were needed (in other words, participants still perceived the stimuli as very painful under placebo analgesia), participants were classified as non-responders.

Study 3 used an additional fourth criterion, as this study included a direct comparison of the participants' hands in a pain task done after the placebo analgesia induction. Pain ratings were compared for the left (control condition) and right (placebo condition) hand. If the average rating of the right hand stimuli was higher than the one for the left hand, participants were classified as non-responders. We thus used the grouping resulting from each study to arrive at our two groups.

2.3 Measures

2.3.1 Behavioral data

In each study, participants had completed a battery of questionnaires measuring different dispositional psychological constructs. For this study, we selected all questionnaires measuring empathy and prosocial behavior as well as empathy-related psychiatric conditions, psychopathy, and depression (see Table 1 for an overview of all questionnaires and which were included in which study). Importantly, not all studies administered all questionnaires, leading to different sample sizes in our group analyses (see also Table 2 below). We included a total of 25 subscales out of eight self-report questionnaires (22 out of which were empathy-related), which are detailed below.

The Empathy Components Questionnaire (ECQ; Batchelder et al., 2017) measures cognitive and affective empathy components, but divided into ability and drive components. It consists of five subscales, (1) cognitive ability, the skill, capacity or potential in perspective taking and to adopt another's point of view; (2) cognitive drive, the motivated interest or tendency in perspective-taking, i.e., to adopt another's point of view; (3) affective ability, the skill, capacity or potential in recognizing, being sensitive to and sharing others' emotional experiences; (4) affective drive, the motivated interest or tendency in recognizing, being sensitive to and sharing others' emotional experiences; and (5) affective reactivity, the emotional response and reaction to other's emotional experiences. Similar to the QCAE, combining all components yields a cumulative total empathy score. The 27 items are rated on 4-point Likert scales ranging from 'strongly disagree' to 'strongly agree'.

The Empathy for Pain Scale (EPS; Giummarra et al., 2015) assesses empathy specific to seeing another individual in pain. It consists of three subscales: (1) affective distress, feelings of distress, discomfort, fear, avoidance, restlessness, and visceral sensations; (2) empathic concern, feelings of compassion, desire to help, and state empathy; and (3) vicarious pain, feelings of both non-painful and painful vicarious sensations. Across four different scenarios (a person undergoing a surgical procedure, a person who has recently had a surgical procedure, a person being accidentally injured, and a person being physically assaulted), 12 identical items are rated on 5-point Likert scales, ranging from 'strongly disagree' to 'strongly agree'.

The Interpersonal Reactivity Index (IRI; Davis et al., 1980; Davis, 1983) is one of the most frequently used self-report questionnaires to measure empathy. It is a multidimensional measure assessing cognitive and affective empathy across four subscales: (1) Fantasy, the tendency to transpose imaginatively into fictional situations; (2) perspective taking, the tendency to adopt the perspectives of others; (3) empathic concern, the tendency to have feelings of warmth, compassion, and concern for others; and (4) personal distress, the tendency to have feelings of discomfort and anxiety when witnessing others' negative experiences. The subscales fantasy and perspective taking aim to measure cognitive aspects, whereas the subscales empathic concern and personal distress aim to measure affective aspects of empathy. Each subscale consists of seven items (28 items in total) that

are each rated on a 5-point Likert scale ranging from ‘does not describe me well’ to ‘describes me very well’.

The Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011) also provides subscales for cognitive and affective empathy, but draws a more clear distinction between these two components. Cognitive empathy is measured using the subscales (1) perspective taking,the tendency to put oneself in another person’s shoes to see things from his or her perspective and (2) online simulation, the effortful attempt to put oneself in another person’s position by imagining what this other person is feeling. Affective empathy is measured using three subscales: (3) Emotional contagion, the automatic tendency to mirror the feelings of others (4) proximal responsivity, the affective response when witnessing the mood of others in a close social context; and (5) peripheral responsivity, similar to proximal responsivity but in a detached context. Combining all five subcomponents provides a cumulative total empathy score. All 31 items are rated on 4-point Likert scales with response options ranging from ‘strongly disagree’ to ‘strongly agree’.

The Vicarious Pain Questionnaire (VPQ; Grice-Jackson et al., 2017) consists of 16 ten-second videos of people going through painful experiences (such as sports injuries or getting injections; for all video stimuli see https://www.youtube.com/channel/UCT8goTgWGRsu14NjVaPCSGw/videos) and measures the extent to which participants experience physical sensations of unpleasantness and/or pain when watching these situations. After watching each video, participants are asked to report whether and how intense they experienced pain in their own body on a scale of 1-10, anchored from ‘very mild’ to ‘intense pain’. These two measures are summed up to a total number of pain responses ranging from 0 to 16, and the average pain intensity of those responses. Participants are further prompted to report the localisation of the felt pain and describe their experience using a list of adjectives, including 10 sensory, 10 affective, and 3 cognitive adjectives (see https://youtu.be/iYkVmLfAt6w for more details on the analysis and score calculation). These variables are used to compute two indices regarding localized vs. generalized and sensory vs. affective experiences. Irrespective of whether participants experienced any vicarious pain, the general frequency of such experiences in everyday life is measured on a 10-point Likert scale from ‘hardly ever’ to ‘very regularly’.

The Helping Attitudes Scale (HAS; Nickell, 1998) is a 20-item measure of respondents’ beliefs, feelings, and behaviors associated with helping. Each item is answered on a 5-point Likert scale from ‘strongly disagree’ to ‘strongly agree’ and summed up to one total score.

The Psychopathy subscale of the Short Dark Triad (SD3; (Jones and Paulhus, 2013) measures participants’ socially aversive, psychopathic behavior. The nine items sum up to one total score.

Lastly, the revised Beck Depression Inventory (BDI-II; Kühner et al., 2007) assesses acute depressive symptomatology using 21 items corresponding to symptoms of depressions as listed in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013). Participants rate the acute severance of the respective symptoms on a 4-point scale, which are then summed up to one total score.

In addition, we used data from the Toronto Alexithymia Scale (TAS; Bagby et al., 1994) and the Autism Quotient (AK-k; Freitag et al., 2015) to check for group differences and possible covariates.

2.3.2 Brain data

For Study 1, structural MRI data were obtained at the Medical University of Vienna, using a 3 Tesla Siemens Tim Trio scanner with a 32-channel head coil. T1-weighted scans of the whole brain were acquired with a magnetization-prepared rapid gradient-echo sequence with the following

parameters: TE/TR = 4.21/2,300 ms, 160 sagittal slices, voxel size = 1.0 × 1.0 × 1.1 mm, field of view = 256 mm. In Studies 2 and 3, data were obtained at the University of Vienna's MRI Center, using a 3 Tesla Siemens MAGNETOM Skyra with a 32-channel head coil. T1-weighted scans of the whole brain were acquired with a magnetization-prepared rapid gradient-echo sequence. Parameters in Study 2 were TE/TR = 2.29/2,300 ms, 176 sagittal slices, voxel size = 0.9 x 0.9 x 0.9 mm, field of view = 240 mm; and in Study 3: TE/TR = 2.43/2,300 ms, 208 sagittal slices, voxel size = 0.8 x 0.8 x 0.8 mm, field of view = 240 mm.

2.4 Data analysis

2.4.1 Behavioral data

To analyze the questionnaire data, we performed group comparisons between participants previously classified as placebo responders and non-responders (see 2.6 above). As we had unequal sample sizes (i.e., a higher number of responders compared to non-responders) as well as non-normally distributed data for most questionnaire scales, we used the more robust Welch's *t*-test for all comparisons (Rasch et al., 2011; Delacre et al., 2017). We also report Cohen's *d* effect sizes as well as 95% confidence intervals (*CIs*) for all group comparisons. Differences between groups were considered statistically significant for *p*-values < .05. For all significant results, we additionally report whether the groups differed in terms of age, gender as well as autistic and depressive traits and the results when including any of those variables into the analysis as a covariate (mean-centered separately for each group). Homogeneity analyses showed that the questionnaire scores were comparable between studies (except the localized-generalized dimension of the VPQ ($p < .001$) and the Empathic Concern subscale of the IRI ($p = 0.42$)). We also exploratorily re-ran our analyses in a linear mixed model with study as a random factor (questionnaire score ~ Responder + (1|Study)). We further exploratorily conducted outlier identification using the interquartile range criterion (Barbato et al., 2011) and repeated the behavioral analyses without those outliers, which did not change the overall results.

This study was data-driven , in the sense that the initial data were not collected for the purpose of the present study. However, as we tested many questionnaires measuring empathy as a concept, we addressed the issue of type-I error inflation by additionally reporting a correction for multiple comparisons using the Benjamini-Hochberg adjustment (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001), which controls the false discovery rate (FDR). While the commonly used Bonferroni correction is judged as too conservative, especially if hypotheses are non-orthogonal, the Benjamini-Hochberg adjustment is considered as less stringent and thus resulting in more powerful tests (García, 2004; Streiner, 2015; Lee and Lee, 2018). In this context, the *p*-values of all empathy-related tests (ECQ, EPS, IRI, QCAE, and VPQ, with a total of 22 subscales/tests) were ordered from the smallest to the largest. Each individual *p* value was then compared to its Benjamini-Hochberg critical value, which was calculated as follows: (rank / total number of tests) x FDR (set at 0.05). The largest *p*-value for which $p <$ critical value was classified as significant and all p-values smaller than the largest value were classified as significant as well. For completion, we therefore report and discuss both the uncorrected results of the empathy questionnaires as well as the results after correcting for multiple comparisons. All analyses were conducted in R version 4.1.2.

2.4.2 Brain data

We used FreeSurfer, a well-validated open source software package that uses a surface-based approach to process and analyze structural brain data by creating a three-dimensional model of the cortical surface (versions 6.0.0 and 7.2.0; http://surfer.nmr.mgh.harvard.edu; Rosas et al., 2002; Kuperberg et al., 2003; Fischl, 2012). For preprocessing and analysis, we used the guidelines in the

Freesurfer short course provided by Andrew Jahn (Andy's Brain Book, https://andysbrainbook.readthedocs.io/en/latest/FreeSurfer/FreeSurfer_Introduction.html). We used Freesurfer's fully automated standard processing pipeline for anatomical brain data, using the recon-all function. The pipeline consists of multiple stages (for a more detailed description, see Dale et al., 1999; Fischl et al., 1999; Fischl and Dale, 2000).

In brief, T1-weighted images underwent the following steps: Skull stripping, intensity normalization, gray-white matter segmentation, and topology correction. Reconstruction of cortical surface models resulted in the gray-white boundary surface as well as the pial surface for each cortical hemisphere. Subcortical regions were segmented, the cortex was parcellated according to the Desikan-Killiany atlas (Desikan et al., 2006), and GMV/CSA/CT as measures of brain morphometry were computed. The resulting maps were smoothed using a 10 mm full-width at half-maximum Gaussian kernel. A total of 174 brains were preprocessed, but reconstruction of one structural image failed due to low image quality resulting from excessive movement during scanning. This participant was excluded, and data analysis was performed on the remaining 173 participants. As the preprocessing pipeline is fully automatic and errors can occur in this process, we conducted an extensive manual quality check procedure and followed the steps described in the tutorial by Andrew Jahn (see link above). We checked both the computed pial and white matter surfaces of the 173 anatomical brain scans by scrolling twice through all brain slices from a coronal view, once focusing on the temporal lobes (as this region is very error-prone) and once focusing on the rest of the brain. If the surface lines were not drawn correctly on three or more consecutive slices, we additionally checked this area in the sagittal and axial view. If the same error was visible in the coronal and one other (sagittal or axial) view, again in at least 3 consecutive slices, we deemed the error meaningful and noted it down. In addition to the two surfaces, we also checked for intensity normalization errors in the white matter. These manual checks were conducted by five individuals (Authors HH and MB as well as three research interns). All quality checkers were trained on five brains that were checked by the main experimenters HH and MB to ensure a homogeneous procedure between individuals, with each brain being checked independently by two of the five people. Non-overlapping errors were discussed until a consensus was reached on whether or not to include those errors and a model-preprocessed brain available in Freesurfer was used for comparison, if needed. After this procedure, major errors were manually corrected using the setting of control points, rerunning parts of recon-all until the error was not visible anymore. Analyses were then performed on the corrected structural brain images.

To examine structural differences between placebo responders and non-responders, we computed a whole-brain analysis for both hemispheres and each outcome measure - GMV, CSA and CT. First, all individual reconstructed cortical surfaces were aligned to an average template using the mri_preproc function in FreeSurfer. Then, intergroup comparisons were performed using vertex-by-vertex general linear models by applying the mri_glmfit function, separate for each hemisphere and measure. Cluster-wise correction for multiple comparisons was performed for all results using pre-computed z Monte Carlo simulations within the function mri_glmfit-sim, which we used to calculate cluster-corrected maps for each contrast. A cluster-wise threshold of $p < .05$ was used and the vertex-wise criterion for statistical significance was set at $p < .001$ (two-sided test). Results are reported using the contrasts responders vs. non-responders and vice versa.

As two different scanners were used over the whole dataset, we also explored the effect of scanner by including it as a covariate in our whole-brain GLM. We further explored the role of gender as a covariate (for these additional analyses see Supplement).

3 Results

3.1 Behavioral results

Table 3 presents the uncorrected results for each questionnaire subscale, which revealed significant group differences in three scales, data of which are displayed in Figure 1 and are reported in the order of the highest to the lowest effect size.

First, placebo responders reported to show significantly higher empathic concern specifically related to pain ($M \pm SD = 3.56 \pm 0.71$) than non-responders ($M \pm SD = 2.99 \pm 0.89$; $t(30.65) = 2.62$, $p = .013$, $d = 0.71$, 95% *CI* [-1.01, -0.13]) in the EPS. The EPS was only measured in Study 3. Second, placebo responders self-reported significantly higher prosocial tendencies ($M \pm SD = 78.06 \pm 8.98$) than non-responders ($M \pm SD = 68.64 \pm 9.74$; $t(13.83) = -2.95$, $p = .011$, $d = 1.01$, 95% *CI* [-16.27, -2.57]) in the HAS. Lastly, we observed that placebo responders reported to show significantly lower psychopathic traits ($M \pm SD = 17.90 \pm 4.22$) compared to non-responders ($M \pm SD = 22.64 \pm 4.52$; $t(13.92) = 3.19$, $p = .007$, $d = -1.08$, 95% *CI* [1.55, 7.92]) in the SD3. The HAS and SD3 were only measured in Study 4. In addition, there was a trend for the proximal responsivity subscale of the QCAE, whereby placebo responders reported slightly higher values ($M \pm SD = 11.61 \pm 2.28$) compared to non-responders ($M \pm SD = 10.71 \pm 2.70$; $t(60.83) = -1.95$, $p = .056$, $d = 0.36$, 95% *CI* [-1.82, 0.02]). Importantly, correction for multiple comparisons using the Benjamini-Hochberg adjustment with an FDR of 0.05, revealed no empathy questionnaire p-value smaller than the Benjamini-Hochberg critical value. Thus, no significant group differences in any empathy questionnaire remained after correcting for multiple testing.

With two exceptions, we found no significant group differences regarding age, sex or any autistic/alexithymic/depressive traits in the different questionnaire samples (all p's > .073.259). We did observe a significantly higher amount of females in the IRI responder sample (114 females, 67 males) compared to the non-responder sample (26 females, 30 males) ($p = .041$). We also found a significantly higher mean age in responders ($M \pm SD = 23.94 \pm 2.75$) compared to non-responders ($M \pm SD = 22.38 \pm 2.56$) in the EPS sample ($p = .026$). However, this mean difference of 1.56 years was relatively small and likely negligible. Including gender, age, autistic traits, depressive traits, or alexithymic traits separately as covariates did not change the presented results. The exploratory linear mixed model with study as a random factor revealed the same results as above, but additional group differences in two empathy questionnaires, the Empathic Concern scale of the IRI (p = .045) and the Proximal Responsivity scale of the QCAE (p = .028). In both cases, non-responders had lower scores than responders. However, these two results did not survive the BH correction for multiple comparisons.

– Table 3 here –

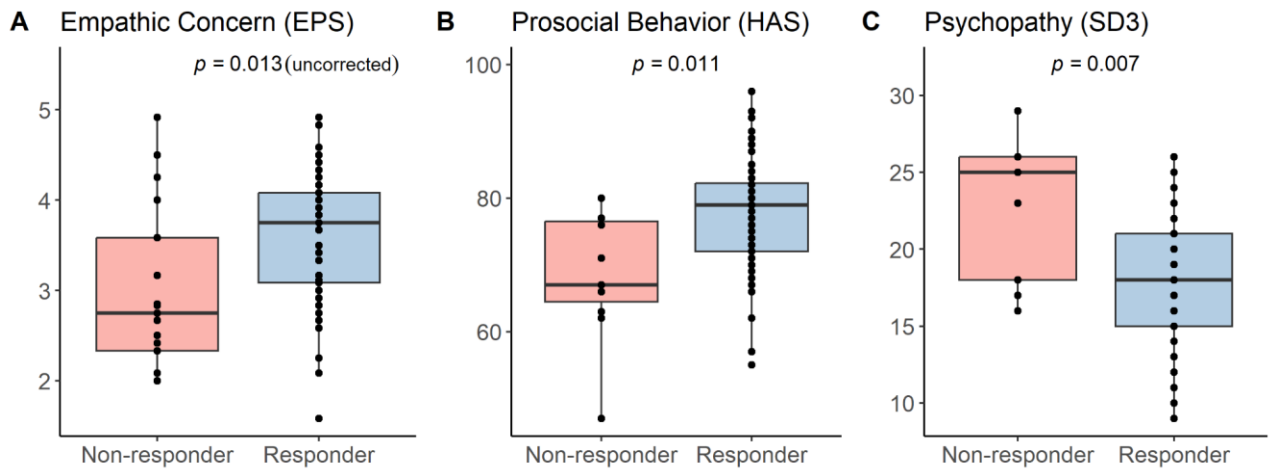

Figure 1. Significant group differences between placebo analgesia responders and non-responders in (A) empathic concern (uncorrected, only measured in Study 3), (B) prosocial behavior (only measured in Study 4), and (C) psychopathy (only measured in Study 4. Correction using the Benjamini-Hochberg adjustment resulted in no significant differences in the empathy subscale (A). Results of all conducted tests including descriptive data and statistics can be found in Table 3.

3.2 Brain results

Next, we investigated structural brain differences on the whole-brain level between the two groups in our aggregated sample of 173 individuals (see Table 4 and Figure 2). The responders and non-responders of the brain data sample did not differ significantly regarding age, gender or any autistic/alexithymic/depressive traits (p's > .100), but they did differ significantly in their total intracranial volume (TIV; $t(70.71) = 2.57$, $p = .012$), their white matter volume ($t(77.59) = 2.33$, $p = .022$), and their cerebrospinal fluid volume ($t(66.85) = 2.19$, $p = .032$). As these differences make the inclusion of them as covariates difficult (Miller and Chapman, 2001), but especially correction for TIV is routinely done in studies assessing structural brain measures such as CSA and GMV that scale with TIV, we therefore report whole-brain results with and without correction for TIV.

In the contrasts non-responder vs. responder, we found increased GMV in left inferior temporal and supramarginal cortices (two clusters) and increased CSA in bilateral middle temporal cortex (three clusters), while there were no differences between the groups regarding CT. We did not find any significant differences in any of the opposite contrasts (responder > non-responder) in any of the three measures. Importantly, when re-running the whole-brain GLMs using TIV as a covariate, no result remained significant at the chosen cluster-wise correction level. Exploratory analyses including scanner or gender as a covariate are reported in the Supplement.

– Table 4 here –

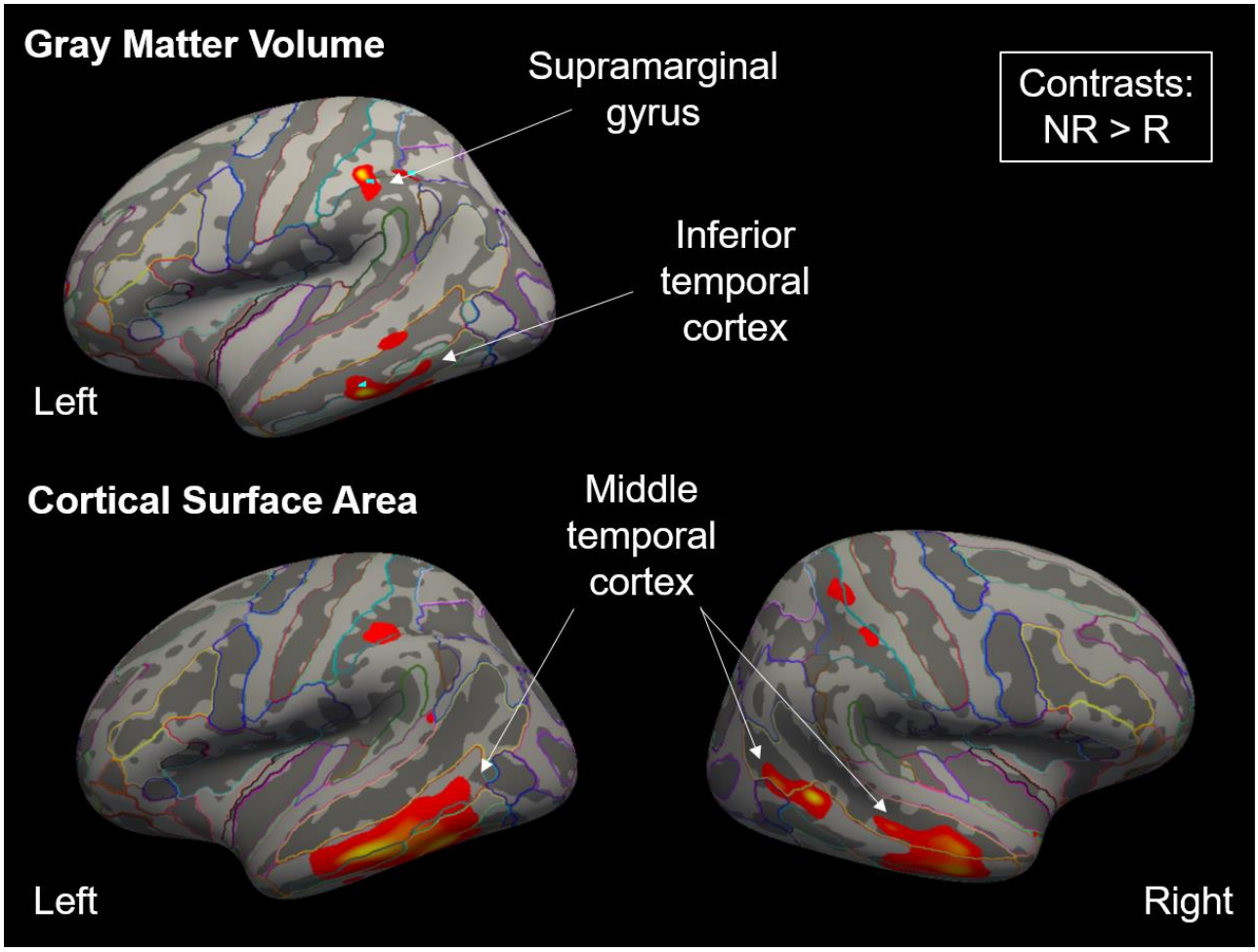

Figure 2. Significant clusters for gray matter volume and cortical surface area when comparing placebo analgesia non-responders (NR) to responders (R), displayed as a heatmap. Clusters were significant for NR > R, while the opposite contrasts did not show significant differences. Colorful outlines show parcellation using the Destrieux atlas. When re-running the whole-brain GLMs using total intracranial volume as a covariate, no result remained significant at the chosen cluster-wise correction level.

4 Discussion

The aim of the present study was to investigate potential differences in psychological traits and brain structure between placebo analgesia responders and non-responders, using a mixed confirmatory and exploratory approach and focusing on social psychological factors. Prior research on the link between psychological characteristics and placebo responding has found few reliable results. In order to provide more clarity and robust results, the current study intended to overcome the limitations of earlier studies by using a larger sample size and stable context factors. Our findings revealed significant differences in empathy-related trait constructs, with placebo responders reporting increased prosocial tendencies, and lower psychopathic traits compared to non-responders. We also observed higher pain-related empathic concern in responders vs. non-responders, although this did not withstand correction for multiple comparisons. Additionally, the neuroimaging data demonstrated smaller gray matter volume and smaller cortical surface area in placebo analgesia responders in supramarginal gyrus, as well as inferior and middle temporal gyrus.

Regarding the behavioral data, placebo responders self-report indicated they display more prosocial and less psychopathic traits, compared to non-responders. Furthermore, responders reported higher empathic concern for the pain of others (uncorrected), while we did not find such group differences in our other measures of general empathy. Interestingly, the pain-specific empathic responses were related to feeling compassion and the need to help others. Even though, this result was non-significant after correcting for multiple comparisons, our results are in line with previous literature on the relationship between these concepts, showing a positive relationship between compassion and prosociality (Karnaze et al., 2022), as well as a negative relationship between social emotions or behavior and psychopathy (Van Dongen, 2020). However, it is important to note that these three results were the only significant ones out of a total of 25 conducted tests. When correcting for 22 empathy-related tests, the empathic concern scale did not reach statistical significance anymore. It is also interesting to note that the other two findings were adjacent but different constructs, prosocial behavior and psychopathic traits. These data thus give hints which questionnaires might be worth further investigation in future studies with larger samples.

The placebo response is a highly complex phenomenon, which is based on a multitude of factors such as person, situation, and/or context. Past research demonstrated that depending on the used paradigm/context, results could vary greatly, and in some cases, even lead to a reversed association between personality traits and the placebo response. For instance, a negative relationship between empathy and the benefit of the placebo treatment was found in a stress paradigm (Darragh et al., 2014), contrasting the frequently reported positive relationship in pain paradigms (Colloca and Benedetti, 2009b; Hunter et al., 2014). Similar observations have also been reported for several other personality traits, such as extraversion and optimism (Darragh et al., 2014). Overall, the link between personality and placebo responding is not straightforward but instead might be modulated by environmental factors such as context and/or situation. Thus, it seems highly probable that, depending on the context, different personality traits become more or less relevant for the placebo response (Whalley et al., 2008; Darragh et al., 2015). In our case, pain-related empathic traits appear to be characteristics in which placebo responders differ from non-responders specifically inside of highly similar pain paradigms. However, this study is the first that included a pain-specific empathy measure so far; therefore, this preliminary assumption needs further investigation. Nevertheless, in order to harness the full potential of the "placebo personality", e.g., to utilize personality traits in order to predict future placebo responding, it might be beneficial to match personality traits and the context of placebo administration.

By systematically including context factors, future studies could further elucidate the question under which circumstances, which personality traits are important for the placebo response. Moreover, it could help identify traits which are stable across different contexts and traits which are changing depending on the context. One such context factor to consider is the relationship quality between healthcare practitioner and patient. It has repeatedly been shown to have a significant influence on the placebo response, whereby factors such as empathy, warmth, communication of positive expectations, and duration of interaction enhance the magnitude of the placebo response (Kaptchuk et al., 2008; Kelley et al., 2009). As none of our included studies measured this variable, this hypothesis remains speculative and future studies could for example measure both patients' and clinicians' empathic abilities and investigate their additive vs. distinct association with social abilities.

In relation to this, our findings also have implications for the translation of artificial lab experiments to real-world settings such as the doctor's office, where individuals may have different motivations (pleasing the experimenter vs. getting help) or need for relief (e.g., no prior complaints vs. symptom reduction). Future studies in this field should thus shift focus towards answering the

question, whether predictors of placebo responsiveness in the lab can be readily transferred to the real world, and if not, how they can be made more useful. The answer to this question may have profound consequences for future research.

In accordance with Coll et al. (2017) we consider empathy as the product of two processes, namely identification of another's emotional state (emotion identification) and sharing the other's emotional state (affect sharing). Even though the amount of empathy an individual has towards another is highly context dependent, there is interindividual variability in how well these two processes are performed. For example, one might demonstrate a lower empathic response due to inadequate emotion identification or due to a limited ability to share the affective state of another. Consequently, in this conceptual frame, some individuals can have higher trait empathy than others by performing better at identifying and sharing another's affective state. Now, our present results imply a positive effect of higher empathy levels, and higher prosocial tendencies in general, on placebo responsiveness. So why could higher prosocial traits lead to enhanced placebo responses? First of all, it needs to be considered that empathy plays an essential role in building relationships. Additionally, the quality of the physician-patient relationship is of particular importance for the treatment outcome. In this respect, a study has demonstrated that the physician's ability to empathize with the patient in combination with the patients' dispositional optimism led to significantly improved health outcomes in chronic pain patients (Cánovas et al., 2018). Furthermore, higher interpersonal trust in the dyadic physician patient relationship has demonstrated similar positive effects (Birkhäuer et al., 2017; Wu et al., 2022). Since earlier lines of research have linked higher dispositional optimism to higher placebo responsiveness (Geers et al., 2010; Morton et al., 2009) we suggest that these positive effects on the placebo response could be explained by the overall emotional valence of the dyadic interaction: traits which have beneficial effects on interpersonal interactions, such as empathy, prosocial traits, optimism, and the absence of obstructive factors (e.g., psychopathy) favor an emotional environment, which facilitates the elicitation of high placebo responses. Moreover, the better the match between personality traits and contextual factors (e.g., pain study paradigm), the higher the chance for a placebo response. In line with this argument, using participants' empathy for other people's pain might be an interesting avenue for predicting their individual hypoalgesic response to a placebo. However, it is important to be cautious with this interpretation as the empathy results were non-significant after correction for multiple comparisons and this should be regarded as preliminary, and in need of replication and extension.

Regarding the brain data, we found smaller gray matter volume and smaller cortical surface area in placebo analgesia responders compared to non-responders in the supramarginal gyrus, as well as the inferior and middle temporal gyrus. Given that placebo responses are the result of complex central nervous system mechanisms, including expectancies and different types of learning processes (Benedetti, 2013), we assumed that measures of brain structure would be able to distinguish placebo responders from non-responders. However, the found regions do not match with the few previous studies, which highlight a positive correlation of gray matter density in brain areas such as the dorsolateral prefrontal cortex, insula, and nucleus accumbens and placebo analgesic effects (Schweinhardt et al., 2009).

Earlier studies suggest that larger cortical surface area is related to higher overall neurocognitive performance whereas smaller cortical surface area is associated with problems in executive functioning (Gutiérrez-Galve et al., 2010; Hartberg et al., 2011; Tamnes et al., 2015; Tadayon et al., 2020). This can be explained by the circumstance that a greater cortical surface area corresponds to a more convoluted sulcal shape with a greater amount of corticocortical connections, which drives local processing. This might increase the performance of brain function and improve cortical computation (Im et al., 2008). Our placebo non-responders demonstrated higher cortical

surface area in several regions compared to placebo responders, which somewhat contradicts our prior assumptions. However, the observed increase in cortical surface area in non-responders could represent an underlying opposing mechanism, which hampers the formation of placebo responses. Thus, higher cortical surface area in these regions could be associated with lower expectancies, empathy (as demonstrated in our results), and overall traits which are detrimental for the placebo response, leading to lower placebo responses. What has to be noted, though, is the absence of typically found regions related to the neural basis of empathy. Previous work highlights negative associations of general empathic traits with gray matter volume in precuneus inferior frontal gyrus, anterior cingulate cortex (Banissy et al., 2012) as well as anterior insula (Li et al., 2020). We, on the other hand, found structural brain differences between responders and non-responders in middle and inferior temporal areas. Thus, the relationship of placebo responsiveness and empathy on the neural level will have to be investigated further in the future.

It is possible that limitations might have influenced our results and therefore, a number of shortcomings should be acknowledged. One caveat of the present study is that our study was exploratory, and that some of the analyses do not withstand correction for multiple comparisons, and the brain results do not hold when correcting for total intracranial volume (TIV). This means that our brain differences could also be the result of inherent brain size differences between the groups. As our two groups differed significantly in their TIV, inclusion of this variable as a covariate becomes difficult. Interestingly, exploratory analyses showed that brain results differed when including gender as a covariate. Moreover, including scanner as a covariate changed the resulting clusters, while identifying similar regions largely located in the temporal cortex. Both of these analyses highlight the need to include them as a covariate in future studies (for gender effects in placebo research see e.g., Enck & Klosterhalfen, 2019). Especially because of this, replication and theoretical extension of our findings is needed. Even when using a pooling approach spanning multiple studies and thus increasing the sample size, like the one used here, power may still be too low to detect these subtle differences using stricter corrections. Future work in this field should therefore focus on using a much larger sample of participants, for example data repositories like the Human Connectome Project (https://www.humanconnectome.org/) and NeuroVault (https://neurovault.org/), or collaborate in multi-lab studies (see e.g. the ManyLabs projects) to investigate this and similar research questions.

Furthermore, we used a convenience sample, pooling data from four of our own lab's placebo studies. An effort towards validation and generalization of the present findings could be achieved by including placebo analgesia studies from multiple labs and even countries. The included studies all come from the field of placebo analgesia research and utilized a similar study design, yet for the present research question, this design could be deficient itself, as it assessed the placebo response only once. For the interpretation of our results, the placebo response was seen as a stable trait of an individual. In other words, subjects who responded to the placebo, are expected to demonstrate this response in the future as well. Several papers have addressed this issue and suggest temporal stability, given that contextual factors remain stable (Whalley et al., 2008; Ashar et al., 2017). Since the temporal stability of the placebo response is an important prerequisite to harness the response for other purposes (e.g., the adjustment of a medical treatment based on an individual's tendency to respond to a placebo), future studies should aim to address this matter in their study designs. Relatedly, we chose to employ a-priori non-responder classification criteria of the included studies as opposed to pain ratings. Future work should investigate direct comparisons of certain criteria and which ones, alone or in combination, best predict placebo (non)responding. Lastly, using previously collected convenience samples meant that not all studies obtained the same trait measures, which is why sample sizes differ in our questionnaire analyses, and especially results from smaller samples should be interpreted with caution until independently replicated.

As evident from the present findings, it remains a challenge to narrow down on a few selected predictive criteria that are able to predict the placebo response. Nevertheless, our study adds valuable information on additional possible social characteristics contributing to the existing evidence. We also propose that placebo responder research needs to move away from the intention to identify single traits which are exclusively related to the magnitude of the placebo response and towards a larger context. Future studies should pursue an interactional approach, by including context-specific factors as well as a large range of psychological and biological characteristics. In this way, context-specific placebo responder patterns can be identified.

Overall, the present study does not provide clear-cut generalizable evidence but draws a heterogenous picture in terms of placebo responder characteristics. Our results broaden the understanding of placebo responder characteristics inside of pain contexts and additionally, we highlight new directions towards social emotions and behavior for exploring differences between responders and non-responders. Importantly, we also provide important hints where future studies could look for biomarkers regarding placebo responsiveness in the brain. While the question "Who responds to placebos?" is still only partly answered, the present study brings us one step closer to answering this crucial question, and paves the way for a better identification of placebo responders in clinical contexts.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

HH: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration, Funding acquisition, Visualization; **MB**: Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization; **FR**: Conceptualization, Methodology, Investigation, Writing - Review & Editing, Supervision, Project administration; **CL**: Conceptualization, Resources, Writing - Review & Editing, Supervision, Funding acquisition.

7 Funding

The present study was financially supported by the uni:docs scholarship of the University of Vienna, the Marietta-Blau scholarship of the Agency for Education and Internationalisation Austria (OeaD) (both to HH), and the Austrian Science Fund (FWF W1262-B29 and P32686 to CL). None of the funders had any role in study design, data collection and analysis, interpretation, writing or decision to publish.

8 Acknowledgments

We thank Markus Rütgen and Isabella Wagner for providing us with the data of their respective studies. We also thank Anna Thallinger, Lisa Schabhüttl and Benedikt Steininger for helping with the manual quality checks of the structural brain data. We thank Patrick Friedrich for insightful conversations regarding the reporting and discussion of our results.

9 References

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders. doi: 10.1176/APPI.BOOKS.9780890425596.

Ashar, Y. K., Chang, L. J., and Wager, T. D. (2017). Brain Mechanisms of the Placebo Effect: An Affective Appraisal Account. Annu. Rev. Clin. Psychol. 13, 73–89. doi: 10.1146/annurev-clinpsy-021815-093015.

Atlas, L. Y., and Wager, T. D. (2014). A meta-analysis of brain mechanisms of placebo analgesia: Consistent findings and unanswered questions. *Handb Exp Pharmacol* 225, 37–69. doi: 10.1007/978-3-662-44519-8_3/COVER.

Bagby, R. M., Parker, J. D. A. & Taylor, G. J. The twenty-item Toronto Alexithymia scale-I. Item selection and cross-validation of the factor structure. J. Psychosom. Res. 38, 23–32 (1994).

Banissy, M. J., Kanai, R., Walsh, V., & Rees, G. (2012). Inter-individual differences in empathy are reflected in human brain structure. Neuroimage, 62(3), 2034-2039.

Barbato, G., Barini, E. M., Genta, G., & Levi, R. (2011). Features and performance of some outlier detection methods. Http://Dx.Doi.Org/10.1080/02664763.2010.545119, 38(10), 2133–2149. https://doi.org/10.1080/02664763.2010.545119

Batchelder, L., Brosnan, M., and Ashwin, C. (2017). *The development and validation of the empathy components questionnaire (ECQ)*. doi: 10.1371/journal.pone.0169185.

Benedetti, F. (2013). Placebo and the new physiology of the doctor-patient relationship. *Physiol Rev* 93, 1207–1246. doi: 10.1152/physrev.00043.2012.

Benedetti, F., Arduino, C., Costa, S., Vighetti, S., Tarenzi, L., Rainero, I., et al. (2006). Loss of expectation-related mechanisms in Alzheimer's disease makes analgesic therapies less effective. *Pain* 121, 133–144. doi: 10.1016/j.pain.2005.12.016.

Benedetti, F., Piedimonte, A., and Frisaldi, E. (2018). How do placebos work? *Eur J Psychotraumatol* 9. doi: 10.1080/20008198.2018.1533370.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300. doi: 10.1111/J.2517-6161.1995.TB02031.X.

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29, 1165–1188. doi: 10.1214/aos/1013699998.

Birkhäuer, J., Gaab, J., Kossowsky, J., Hasler, S., Krummenacher, P., Werner, C., et al. (2017). Trust in the health care professional and health outcome: A meta-analysis. *PLoS One 12*, e0170988. doi: 10.1371/JOURNAL.PONE.0170988.

Cánovas, L., Carrascosa, A. J., García, M., Fernández, M., Calvo, A., Monsalve, V., et al. (2018). Impact of Empathy in the Patient-Doctor Relationship on Chronic Pain Relief and Quality of Life: A Prospective Study in Spanish Pain Clinics. *Pain Medicine 19*, 1304–1314. doi: 10.1093/PM/PNX160.

Colagiuri, B., Schenk, L. A., Kessler, M. D., Dorsey, S. G., and Colloca, L. (2015). The placebo effect: From concepts to genes. *Neuroscience* 307, 171–190. doi: 10.1016/j.neuroscience.2015.08.017.

Coll, M. P., Viding, E., Rütgen, M., Silani, G., Lamm, C., Catmur, C., et al. (2017). Are we really measuring empathy? Proposal for a new measurement framework. *Neurosci Biobehav Rev, 83*, 132–139. doi: 10.1016/J.NEUBIOREV.2017.10.009.

Colloca, L., and Benedetti, F. (2009a). Placebo analgesia induced by social observational learning. *Pain* 144, 28–34. doi: 10.1016/j.pain.2009.01.033.

Colloca, L., and Benedetti, F. (2009b). Placebo analgesia induced by social observational learning. *Pain* 144, 28–34. doi: 10.1016/J.PAIN.2009.01.033.

Corsi, N., and Colloca, L. (2017). Placebo and nocebo effects: The advantage of measuring expectations and psychological factors. *Front Psychol* 8. doi: 10.3389/fpsyg.2017.00308.

Dale, A. . M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and Surface Reconstruction. *Neuroimage* 9, 195–207. doi: 10.1006/nimg.1998.0396.

Darragh, M., Booth, R. J., and Consedine, N. S. (2014). Investigating the "placebo personality" outside the pain paradigm. *J Psychosom Res* 76, 414–421. doi: 10.1016/j.jpsychores.2014.02.011.

Darragh, M., Booth, R. J., and Consedine, N. S. (2015). Who responds to placebos? Considering the "placebo personality" via a transactional model. *Psychol Health Med* 20, 287–295. doi: 10.1080/13548506.2014.936885.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol* 44, 113–126. doi: 10.1037/0022-3514.44.1.113.

Davis, M. (1980). A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog Sel. Doc. Psychol. 10*.

Delacre, M., Lakens, D., and Leys, C. (2017). Why psychologists should by default use welch's t-Test instead of student's t-Test. *International Review of Social Psychology* 30, 92–101. doi: 10.5334/irsp.82.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021.

Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53, 1–15. doi: 10.1016/J.NEUROIMAGE.2010.06.010.

Elsenbruch, S., Kotsis, V., Benson, S., Rosenberger, C., Reidick, D., Schedlowski, M., et al. (2012). Neural mechanisms mediating the effects of expectation in visceral placebo analgesia: An fMRI study in healthy placebo responders and nonresponders. *Pain* 153, 382–390. doi: 10.1016/j.pain.2011.10.036.

Enck, P., Bingel, U., Schedlowski, M., and Rief, W. (2013). The placebo response in medicine: Minimize, maximize or personalize? *Nat Rev Drug Discov* 12, 191–204. doi: 10.1038/nrd3923.

Enck, P., & Klosterhalfen, S. (2019). Does sex/gender play a role in placebo and nocebo effects? Conflicting evidence from clinical trials and experimental studies. Frontiers in Neuroscience, 13, 160.

Fischl, B. (2012). FreeSurfer Authos Manuscript. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021.FreeSurfer.

Fischl, B., and Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* 97, 11050–11055. doi: 10.1073/pnas.200033797.

Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical Surface-Based Analysis. 207, 195–207.

Freitag, C. M. et al. Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ) - die Kurzversion AQ-K. Z. Klin. Psychol. Psychother. 36, 280–289 (2015).

García, L. v. (2004). Escaping the Bonferroni iron claw in ecological studies. *Oikos* 105, 657–663. doi: 10.1111/J.0030-1299.2004.13046.X.

Geers, A. L., Helfer, S. G., Kosbab, K., Weiland, P. E., and Landry, S. J. (2005). Reconsidering the role of personality in placebo effects: Dispositional optimism, situational expectations, and the placebo response. *J Psychosom Res* 58, 121–127. doi: 10.1016/j.jpsychores.2004.08.011.

Geers, A. L., Wellman, J. A., Fowler, S. L., Helfer, S. G., and France, C. R. (2010). Dispositional optimism predicts placebo analgesia. *Journal of Pain* 11, 1165–1171. doi: 10.1016/j.jpain.2010.02.014.

Giummarra, M. J., Fitzgibbon, B. M., Georgiou-Karistianis, N., Beukelman, M., Verdejo-Garcia, A., Blumberg, Z., et al. (2015). Affective, sensory and empathic sharing of another's pain: The

763 Empathy for Pain Scale. *European Journal of Pain (United Kingdom)* 19, 807–816. doi:
764 10.1002/ejp.607.
765 Grice-Jackson, T., Critchley, H. D., Banissy, M. J., and Ward, J. (2017). Common and distinct
766 neural mechanisms associated with the conscious experience of vicarious pain. *Cortex* 94, 152–
767 163. doi: 10.1016/J.CORTEX.2017.06.015.
768 Gutiérrez-Galve, L., Wheeler-Kingshott, C. A. M., Altmann, D. R., Price, G., Chu, E. M., Leeson,
769 V. C., et al. (2010). Changes in the Frontotemporal Cortex and Cognitive Correlates in First-
770 Episode Psychosis. *Biol Psychiatry* 68, 51–60. doi: 10.1016/J.BIOPSYCH.2010.03.019.
771 Hartberg, C. B., Sundet, K., Rimol, L. M., Haukvik, U. K., Lange, E. H., Nesvåg, R., et al. (2011).
772 Brain Cortical Thickness and Surface Area Correlates of Neurocognitive Performance in
773 Patients with Schizophrenia, Bipolar Disorder, and Healthy Adults. *Journal of the International*
774 *Neuropsychological Society* 17, 1080–1093. doi: 10.1017/S1355617711001081.
775 Hartmann, H., Forbes, P. A. G., Rütgen, M., and Lamm, C. (2022). Placebo Analgesia Reduces
776 Costly Prosocial Helping to Lower Another Person's Pain. *Psychol Sci*, 09567976211197. doi:
777 10.1177/09567976221119727/ASSET/IMAGES/LARGE/10.1177_09567976221119727-
778 FIG2.JPEG.
779 Hartmann, H., Rütgen, M., Riva, F., and Lamm, C. (2021). Another's pain in my brain: No evidence
780 that placebo analgesia affects the sensory-discriminative component in empathy for pain.
781 *Neuroimage* 224, 117397. doi: 10.1016/J.NEUROIMAGE.2020.117397.
782 Horing, B., Weimer, K., Muth, E. R., and Enck, P. (2014). Prediction of placebo responses: A
783 systematic review of the literature. *Front Psychol* 5, 1–10. doi: 10.3389/fpsyg.2014.01079.
784 Hunter, T., Siess, F., and Colloca, L. (2014). Socially induced placebo analgesia: A comparison of a
785 pre-recorded versus live face-to-face observation. *European Journal of Pain (United Kingdom)*
786 18, 914–922. doi: 10.1002/j.1532-2149.2013.00436.x.
787 Im, K., Lee, J. M., Lyttelton, O., Kim, S. H., Evans, A. C., and Kim, S. I. (2008). Brain Size and
788 Cortical Structure in the Adult Human Brain. *Cerebral Cortex* 18, 2181–2191. doi:
789 10.1093/CERCOR/BHM244.
790 Jakšic, N., Aukst-Margetic, B., and Jakovljevic, M. (2013). Does personality play a relevant role in
791 the placebo effect? *Psychiatr Danub* 25, 17–23.
792 Jones, D. N., and Paulhus, D. L. (2013). Introducing the Short Dark Triad (SD3).
793 *http://dx.doi.org/10.1177/1073191113514105* 21, 28–41. doi: 10.1177/1073191113514105.
794 Kang, H., Miksche, M. S., and Ellingsen, D.-M. (2022). The association between personality traits
795 and placebo effects: A preregistered systematic review and meta-analysis. Pain, 10-1097.
796 Kaptchuk, T. J., Kelley, J. M., Deykin, A., Wayne, P. M., Lasagna, L. C., Epstein, I. O., et al.
797 (2008). Do "placebo responders" exist? *Contemp Clin Trials* 29, 587–595. doi:
798 10.1016/j.cct.2008.02.002.
799 Karnaze, M. M., Bellettiere, J., & Bloss, C. S. (2022). Association of compassion and empathy with
800 prosocial health behaviors and attitudes in a pandemic. *PloS one*, *17*(7), e0271829.
801 Kelley, J. M., Lembo, A. J., Ablon, J. S., Villanueva, J. J., Conboy, L. A., Levy, R., et al. (2009).
802 Patient and practitioner influences on the placebo effect in irritable bowel syndrome.
803 *Psychosom Med* 71, 789–797. doi: 10.1097/PSY.0b013e3181acee12.
804 Kern, A., Kramm, C., Witt, C. M., and Barth, J. (2020). The influence of personality traits on the
805 placebo/nocebo response: A systematic review. *J Psychosom Res* 128, 109866. doi:
806 10.1016/j.jpsychores.2019.109866.
807 Koban, L., Ruzic, L., and Wager, T. D. (2013). *Brain Predictors of Individual Differences in*
808 *Placebo Responding*. Elsevier doi: 10.1016/B978-0-12-397928-5.00010-6.
809 Kühner, C., Bürger, C., Keller, F., and Hautzinger, M. (2007). [Reliability and validity of the
810 Revised Beck Depression Inventory (BDI-II). Results from German samples]. *Nervenarzt* 78,
811 651–656. doi: 10.1007/S00115-006-2098-7.

Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., et al. (2003). Regionally localized thinning of the cerebral cortex in schizophrenia. *Arch Gen Psychiatry* 60, 878–888. doi: 10.1001/archpsyc.60.9.878.

Lee, S., and Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol* 71, 353–360. doi: 10.4097/kja.d.18.00242.

Li, Y., Zhang, T., Li, W., Zhang, J., Jin, Z., & Li, L. (2020). Linking brain structure and activation in anterior insula cortex to explain the trait empathy for pain. Human brain mapping, 41(4), 1030-1042.

Liu, J., Mu, J., Liu, Q., Dun, W., Zhang, M., and Tian, J. (2017). Brain structural properties predict psychologically mediated hypoalgesia in an 8-week sham acupuncture treatment for migraine. *Hum Brain Mapp* 38, 4386–4397. doi: 10.1002/hbm.23667.

Miller, G. A., and Chapman, J. P. (2001). Misunderstanding Analysis of Covariance. *J Abnorm Psychol* 110, 40–48. doi: 10.1037/0021-843X.110.1.40.

Morton, D. L., Watson, A., El-Deredy, W., and Jones, A. K. P. (2009). Reproducibility of placebo analgesia: Effect of dispositional optimism. *Pain* 146, 194–198. doi: 10.1016/j.pain.2009.07.026.

Nickell, G. S. (1998). The Helping Attitude Scale. *Annual Convention of the American Psychological Association* 56563, 1–10.

Peciña, M., Azhar, H., Love, T. M., Lu, T., Fredrickson, B. L., Stohler, C. S., et al. (2013). Personality trait predictors of placebo analgesia and neurobiological correlates. *Neuropsychopharmacology* 38, 639–646. doi: 10.1038/npp.2012.227.

Peciña, M., and Zubieta, J. K. (2015). Molecular mechanisms of placebo responses in humans. *Mol Psychiatry* 20, 416–423. doi: 10.1038/mp.2014.164.

Peiris, N., Blasini, M., Wright, T., and Colloca, L. (2018). The Placebo Phenomenon: A Narrow Focus on Psychological Models. *Perspect Biol Med* 61, 388–400. doi: 10.1353/PBM.2018.0051.

Petrovic, P., Kalso, E., Petersson, K. M., and Ingvar, M. (2002). Placebo and opioid analgesia - Imaging a shared neuronal network. *Science (1979)* 295, 1737–1740. doi: 10.1126/science.1067176.

Rasch, D., Kubinger, K. D., and Moder, K. (2011). The two-sample t test: Pre-testing its assumptions does not pay off. *Statistical Papers* 52, 219–231. doi: 10.1007/s00362-009-0224-x.

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., and Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *J Pers Assess* 93, 84–95. doi: 10.1080/00223891.2010.528484.

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., et al. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology* 58, 695–701. doi: 10.1212/WNL.58.5.695.

Rütgen, M., Seidel, E. M., Silani, G., Riečanský, I., Hummer, A., Windischberger, C., et al. (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proc Natl Acad Sci U S A* 112, E5638–E5646. doi: 10.1073/pnas.1511269112.

Schedlowski, M., Enck, P., Rief, W., and Bingel, U. (2015). Neuro-bio-behavioral mechanisms of placebo and nocebo responses: Implications for clinical trials and clinical practice. *Pharmacol Rev* 67, 697–730. doi: 10.1124/pr.114.009423.

Scheier, M. E., and Carver, C. S. (1987). Dispositional Optimism and Physical Well-Being: The Influence of Generalized Outcome Expectancies on Health. *J Pers* 55, 169–210. doi: 10.1111/j.1467-6494.1987.tb00434.x.

Schweinhardt, P., Seminowicz, D. A., Jaeger, E., Duncan, G. H., and Bushnell, M. C. (2009). The anatomy of the mesolimbic reward System: A link between personality and the placebo

analgesic response. *Journal of Neuroscience* 29, 4882–4887. doi: 10.1523/JNEUROSCI.5634-08.2009.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2012). A 21 Word Solution. *SSRN Electronic Journal*. doi: 10.2139/SSRN.2160588.

Singer, T., and Lamm, C. (2009). The Social Neuroscience of Empathy. *Ann N Y Acad Sci* 1156, 81–96. doi: 10.1111/J.1749-6632.2009.04418.X.

Stahl, S. M., and Greenberg, G. D. (2019). Placebo response rate is ruining drug development in psychiatry: why is this happening and what can we do about it? *Acta Psychiatr Scand* 139, 105–107. doi: 10.1111/acps.13000.

Staskin, D. R., Michel, M. C., Sun, F., Guan, Z., and Morrow, J. D. (2012). The effect of elective sham dose escalation on the placebo response during an antimuscarinic trial for overactive bladder symptoms. *Journal of Urology* 187, 1721–1726. doi: 10.1016/j.juro.2011.12.052.

Stevens, F., & Taber, K. (2021). The neuroscience of empathy and compassion in pro-social behavior. *Neuropsychologia, 159*, 107925.

Streiner, D. L. (2015). Best (but oft-forgotten) practices: The multiple problems of multiplicity-whether and how to correct for many statistical tests. *American Journal of Clinical Nutrition* 102, 721–728. doi: 10.3945/ajcn.115.113548.

Tadayon, E., Pascual-Leone, A., and Santarnecchi, E. (2020). Differential Contribution of Cortical Thickness, Surface Area, and Gyrification to Fluid and Crystallized Intelligence. *Cerebral Cortex* 30, 215–225. doi: 10.1093/CERCOR/BHZ082.

Tamnes, C. K., Zeller, B., Amlien, I. K., Kanellopoulos, A., Andersson, S., Due-Tønnessen, P., et al. (2015). Cortical surface area and thickness in adult survivors of pediatric acute lymphoblastic leukemia. *Pediatr Blood Cancer* 62, 1027–1034. doi: 10.1002/PBC.25386.

Vachon-Presseau, E., Berger, S. E., Abdullah, T. B., Huang, L., Cecchi, G. A., Griffith, J. W., et al. (2018). Brain and psychological determinants of placebo pill response in chronic pain patients. *Nat Commun* 9. doi: 10.1038/s41467-018-05859-1.

VanderWeele, T. J., and Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: is the Bonferroni correction really so bad? *Am J Epidemiol* 188, 617–618. doi: 10.1093/AJE/KWY250.

Van Dongen, J. D. (2020). The empathic brain of psychopaths: From social science to neuroscience in empathy. *Frontiers in Psychology*, *11*, 695.

Wager, T. D., and Atlas, L. Y. (2015). The neuroscience of placebo effects: Connecting context, learning and health. *Nat Rev Neurosci* 16, 403–418. doi: 10.1038/nrn3976.

Whalley, B., Hyland, M. E., and Kirsch, I. (2008). Consistency of the placebo effect. *J Psychosom Res* 64, 537–541. doi: 10.1016/j.jpsychores.2007.11.007.

Wu, Q., Jin, Z., and Wang, P. (2022). The relationship between the physician-patient relationship, physician empathy, and patient trust. *J Gen Intern Med, 37*, 1388–1393. doi: 10.1007/S11606-021-07008-9/FIGURES/2.

Zunhammer, M., Spisák, T., Wager, T. D., Bingel, U., Atlas, L., Benedetti, F., et al. (2021). Meta-analysis of neural systems underlying placebo analgesia from individual participant fMRI data. *Nature Communications 2021 12:1* 12, 1–11. doi: 10.1038/s41467-021-21179-3.

10 Data Availability Statement

The dataset, analysis code, and Benjamini-Hochberg correction for this study can be found in the corresponding OSF project (https://osf.io/nwgdj/?view_only=d3d8e54e20f4453fba0ea10b878bb0f3).

907 **11 Tables**

Table 1

Descriptive overview of study samples and measures.

	Sample			Questionnaires								
Study	*n*	R	NR	IRI	QCAE	BDI	VPQ	ECQ	EPS	SD3	HAS	T1
1	63	49	14	x								x
2	37	27	10	x	x	x		x				x
3	74	53	21	x	x	x	x	x	x			x
4	63	52	11	x	x	x	x			x	x	
Total	237	181	56	237	174	174	137	111	74	63	63	173[a]

Note. R = placebo analgesia responder; NR = placebo analgesia non-responder; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy; BDI = Beck Depression Inventory; ECQ = Empathy Components Questionnaire; EPS = Empathy for Pain Scale; VPQ = Vicarious Pain Questionnaire; SD3 = Short Dark Triad; HAS = Helping Attitudes Scale; T1 = T1-weighted structural image. [a]After exclusion of one participant due to excessive movement.

908
909

910

Table 2

Exclusion criteria of each study.

	Study			
Exclusion Criterion	1	2	3	4
Non-German speaker	x		x	x
Years of age < 18 or > 35		x	x	x
Past or present enrollment in psychology, pharmaceutics, or medicine studies		x[a]	x	x[b]
Participation in similar studies (e.g., pain, empathy, or placebo studies)		x		x
Non right-handedness	x	x	x[c]	x
Past or present medical conditions interfering with current pain sensitivity		x	x	x
Neurological or psychiatric conditions	x	x	x	x
Past or present self-injurious behavior			x	
Past or present alcohol misuse	x	x	x	x
Past or present drug misuse	x	x	x	x
Psychopharmacological medication within the last 3 months (apart from oral contraceptives)			x	x
Lactose intolerance (due to the placebo pill containing lactose)		x		
Non MR-safe (e.g., implants, claustrophobia, or pregnancy)	x	x	x	

Note. [a] This study included bachelor psychology students in their first semester. [b] This study only included enrolled students. [c] This study additionally excluded people with a weakness to distinguish left from right.

911

Table 3

Group comparisons of personality/questionnaire scores of placebo responders and non-responders.

	R			NR						95% CI			95% CI	
Questionnaire Subscale	*n*	*M*	*SD*	*n*	*M*	*SD*	*t* (*df*)	*p*	BH	*LL*	*UL*	Cohen's *d*	*LL*	*UL*
ECQ[2,3]														
Affective Ability	80	14,96	2,46	31	14,23	2,84	-1.27 (48.42)	.209	,011	-1,90	0,43	0,28	-0,71	0,14
Affective Drive	80	13,22	1,82	31	12,9	1,3	-1.04 (76.22)	.302	,020	-0,94	0,30	0,20	-0,60	0,20
Affective Reactivity	80	21,4	2,88	31	21	3,79	-0.53 (44.16)	.598	,032	-1,92	1,12	0,12	-0,61	0,30
Cognitive Ability	80	18,73	2,41	31	18,06	2,52	-1.26 (52.66)	.215	,014	-1,72	0,40	0,27	-0,67	0,15
Cognitive Drive	80	16,09	2,19	31	16,03	2,01	-0.13 (59.36)	.900	,045	-0,93	0,82	0,03	-0,45	0,41
EPS[3]														
Affective Distress	53	2,35	0,72	21	2,27	0,77	-0.45 (34.48)	.656	,036	-0,48	0,31	0,12	-0,69	0,42
Empathic Concern	53	3,56	0,71	21	2,99	0,89	-2.62 (30.65)	.013	,002	-1,01	-0,13	0,71	-1,47	-0,17
Vicarious Pain	53	1,95	0,87	21	1,8	0,75	-0.75 (42.39)	.459	,027	-0,56	0,26	0,19	-0,74	0,27
IRI[1-4]														
Empathic Concern	181	19,8	4,73	56	18,41	5,08	-1.82 (86.52)	.073	,007	-2,91	0,13	0,28	-0,60	0,00
Fantasy	181	18,81	5,8	56	18,82	5,23	0.01 (100.50)	.991	,050	-1,62	1,64	0,00	-0,31	0,32
Personal Distress	181	11,45	4,85	56	10,68	5,01	-1.02 (89.21)	.311	,023	-2,29	0,74	0,16	-0,45	0,14
Perspective Taking	181	19,12	4,85	56	18,21	4,74	-1.24 (93.46)	.219	,018	-2,35	0,55	0,19	-0,48	0,12
QCAE[2-4]														
Emotional Contagion	132	10,7	2,4	42	10,43	2,76	-0.58 (61.89)	.563	,030	-1,22	0,67	0,11	-0,48	0,24
Online Simulation	132	27,23	4,05	42	26,24	4,02	-1.40 (69.60)	.167	,009	-2,42	0,43	0,25	-0,63	0,10
Peripheral Responsivity	132	10,77	2,77	42	11,21	2,52	0.98 (74.90)	.330	,025	-0,46	1,36	-0,17	-0,15	0,53
Proximal Responsivity	132	11,61	2,28	42	10,71	2,7	-1.95 (60.83)	.056	,005	-1,82	0,02	0,36	-0,76	-0,01
Perspective Taking	132	30,77	4,25	42	30,36	5,11	-0.48 (60.11)	,635	,034	-2,16	1,33	0,09	-0,45	0,26
VPQ[3,4]														
Intensity	86	3,56	1,95	24	3,6	2,05	0.10 (35.44)	,921	,048	-0,90	1,00	-0,02	-0,43	0,45
Localized-generalized	104	-2,32	5,05	31	-2,1	4,74	0.22 (52.04)	,824	,043	-1,75	2,20	-0,05	-0,36	0,49
Regularity	82	4,06	1,85	25	3,44	2,26	-1.25 (34.48)	,219	,016	-1,63	0,39	0,30	-0,84	0,16
Sensory-affective	105	-0,73	10,65	32	-1,31	8,67	-0.31 (62.14)	,756	,039	-4,28	3,12	0,06	-0,46	0,31
Total pain response	105	5,01	4,51	32	4,72	4,81	-0.30 (48.79)	,763	,041	-2,22	1,64	0,06	-0,52	0.37
HAS[4]														
Total Score	52	78,06	8,98	11	68,64	9,74	-2.95 (13.83)	.011	---	-16,27	-2,57	1,01	-1,83	-0,41
SD3[4]														
Psychopathy	52	17,9	4,22	11	22,64	4,52	3.19 (13.92)	,007	---	1,55	7,92	-1,08	0,45	2,01
BDI[2-4]														

Total Score	132	6,45	6,29	42	6,69	5,72	0.23 (75.14)	.821	---	-1,83	2,31	-0,04	-0,32	0,39

Note. R = placebo analgesia responder; NR = placebo analgesia non-responder; IRI = Interpersonal Reactivity Index; QCAE = Questionnaire of Cognitive and Affective Empathy; BDI = Beck Depression Inventory; ECQ = Empathy Components Questionnaire; EPS = Empathy for Pain Scale; VPQ = Vicarious Pain Questionnaire; SD3 = Short Dark Triad; HAS = Helping Attitudes Scale; CI = confidence interval; BH = Benjamini-Hochberg critical value. Superscript numbers refer to the studies where a certain questionnaire was used.

913

914

Table 4

Structural brain differences for the contrast placebo analgesia non-responders > responders without covariates.

Measure and brain region	h	VtxMax	size	*x*	*y*	*z*	*p*	CI (*p*)
GMV								
Inferior temporal	L	120836	338.40	-54.0	-33.6	-16.9	.003	[.002, .004]
Supramarginal	L	123297	249.71	-44.3	-32.2	34.6	.017	[.013, .017]
CSA								
Middle temporal	R	22997	1259.90	61.3	-14.3	-21.9	< .001	[< .0001, .0004]
Middle temporal	R	91021	452.16	57.9	-57.4	5.0	.010	[.008, .012]
Middle temporal	L	49559	2389.60	-59.7	-28.6	-16.8	< .001	[< .0001, .0004]
CT								
No significant clusters								

Note. Significant clusters separate for gray matter volume (GMV), cortical surface area (CSA) and cortical thickness (CT), including hemisphere h, vertex number at maximum (VtxMax), cluster surface area in mm^2 (size), MNI coordinates x, y, z, cluster-wise *p*-value (threshold of $p < .05$, vertex-wise criterion for statistical significance at $p < .001$, two-sided) and the 90% confidence intervals (CI) of that *p*-value.

915