

Cognition and Emotion

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pcem20

Illustrating the pathway from affect to somatic symptom: the Affective Picture Paradigm*

Tara M. Petzke, Kathrin Weber, Omer Van den Bergh & Michael Witthöft

To cite this article: Tara M. Petzke, Kathrin Weber, Omer Van den Bergh & Michael Witthöft (27 Feb 2024): Illustrating the pathway from affect to somatic symptom: the Affective Picture Paradigm*, Cognition and Emotion, DOI: 10.1080/02699931.2024.2319273

To link to this article: https://doi.org/10.1080/02699931.2024.2319273

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 27 Feb 2024.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=pcem20

COGNITION AND EMOTION
https://doi.org/10.1080/02699931.2024.2319273

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Illustrating the pathway from affect to somatic symptom: the Affective Picture Paradigm*

Tara M. Petzke [a], Kathrin Weber[a], Omer Van den Bergh [b] and Michael Witthöft [a]

[a]Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, Johannes-Gutenberg-University Mainz, Mainz, Germany; [b]Health Psychology, University of Leuven, Leuven, Belgium

ABSTRACT

High levels of somatic symptom distress represent a core component of both mental and physical illness. The exact aetiology and pathogenesis of this transdiagnostic phenomenon remain largely unknown. The Affective Picture Paradigm (APP) represents an innovative experimental paradigm to study somatic symptom distress. Based on the HiTOP framework and a population-based sampling approach, associations between facets of somatic symptom distress and symptoms induced by the APP were explored in two studies ($N_1 = 201$; $N_2 = 254$) using structural equation bi-factor models. Results showed that the APP effect was significantly positively correlated with general somatic symptom distress (PHQ-15, HiTOP), cardio-respiratory symptoms (PHQ-15), as well as difficulties identifying feelings. In conclusion, negative affective cues in the APP can elicit somatic symptoms, particularly in people with higher levels of somatic symptom distress. Difficulties identifying feelings might contribute to this phenomenon. Results are compatible with a predictive processing account of somatic symptom perception.

ARTICLE HISTORY
Received 25 April 2023
Revised 5 October 2023
Accepted 9 October 2023

KEYWORDS
Interoception; affect; symptom perception; Alexithymia; functional disorders

Introduction

Elle Woods, protagonist of *Legally Blonde,* devours a box of chocolates after her boyfriend dumps her. The pain in Bryan Mills' face is palpable when he hears his daughter was abducted in *Taken*. Marta from the mystery thriller *Knives Out* has to vomit every time she tells a lie. These reactions make sense to us: negative emotions seem to be capable of triggering strong bodily sensations and somatic reactions.

One research area at the crossroads of mind and body is interoception. Interoception is defined as the perception and cognitive processing of bodily signals (Khalsa et al., 2018). Two facets are usually distinguished: interoceptive accuracy and interoceptive attention (Murphy et al., 2019). Accuracy comprises how well one perceives bodily states, while interoceptive attention covers the degree to which persons attend to their body. A 2*2 model has been proposed: both facets can be measured using questionnaires and behavioural paradigms, such as experience sampling for interoceptive attention (Csikszentmihalyi & Larson, 2014) and heartbeat perception tasks for interoceptive accuracy (Murphy et al., 2019).

Impaired interoception has been linked to a variety of psychopathological conditions, such as eating disorders (Jenkinson et al., 2018), autism (Garfinkel et al., 2016) and panic disorders (Ehlers, 1993). A currently predominant model on how bodily signals are integrated is the Predictive Coding Model. Predictive coding is an active Bayesian inference process to symptom perception

CONTACT Tara M. Petzke tpetzke@uni-mainz.de Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, Johannes-Gutenberg-University Mainz, Wallstraße 3, Mainz 55122, Germany

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699931.2024.2319273.

*Our preregistration is available at https://doi.org/10.17605/OSF.IO/43PR9. Our study data and analysis code are also available in this repository.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

(Friston, 2010; Van den Bergh et al., 2017). It describes the interaction between priors (top-down processes representing implicit expectations of bodily sensations) and sensory input (or observation) which eventually leads to a posterior model, comprising the resulting perception. Depending on the relative reliability (i.e. precision) of the prior and the sensory input, the posterior will be closer to the sensory input or closer to the prior. In persons with distorted interception and a dominance of the top-down prior, the integration of the bodily signal is therefore more likely to reflect the expectation than the actual state of the observation (Van den Bergh et al., 2017).

One of the most prominent disorders where alterations in interoception seem to be at play are somatoform and related disorders (Bonaz et al., 2021). The prevalence of somatoform disorders is approximately 4.9% (Wittchen et al., 2011). On a symptom level, somatic symptoms unrelated to objectifiable physiological dysfunction make up two thirds of all symptoms reported in primary care (Steinbrecher et al., 2011). These symptoms therefore pose a huge burden to the primary care system, leaving both practitioners and patients unsure about the cause of the symptoms. It further leads to artificial diagnostic categories defined by different (secondary care) specialities (Rasmussen & Rø, 2018; Sirri et al., 2017; World Health Organisation, 2004).

A dimensional approach to psychopathology in general (including "somatoform" psychopathology) is provided by the Hierarchical Taxonomy of Psychopathology (Kotov et al., 2017). Below a general factor of psychopathology at the highest level, the HiTOP framework currently assumes three overarching superspectra, namely the emotional dysfunction, the externalising, and the psychosis superspectrum (Kotov et al., 2020; Krueger et al., 2021; Watson et al., 2022). On the next lower level, the emotional dysfunction superspectrum contains somatoform and internalising psychopathology as overlapping but distinct spectra (Hartmann et al., 2022; Simms et al., 2022; Watson et al., 2022). So recognising persistent somatic symptom distress as part of an emotional dysfunction superspectrum (closely connected to internalising psychopathology) could improve diagnosis and therapy and encourage research into comorbidities and correlations with other (sub)-clinical features.

Concerning the exact psychometric structure of somatic symptom distress, bi-factorial structural models have been found to represent persistent somatic symptom distress well: there seems to be a general overarching affective-motivational factor of symptom perception (motivation to attend to the symptoms and overreactive negative evaluation of these) as well as lower-order symptom clusters, representing specific sensory components of symptom perception (Walentynowicz et al., 2018; Witthöft, Hiller, Loch, & Jasper, 2013, 2016, 2020). Interestingly, while the affective-motivational factor is enhanced in persons with persistent somatic symptoms, there seems to be a less detailed somatosensory processing (Van den Bergh et al., 2021). This is also consistent with the predictive coding model: the bottom-up process appears to be toned down, resulting in less precise sensory input and low interoceptive accuracy, whereas the top-down process dominates the integration efforts due to highly precise (symptom) priors.

Negative affective states and emotions appear to be strongly related to somatic symptom distress. Most researchers agree that emotions rely in part on interoception, as certain bodily processes are signalled to the brain to create an emotional experience, in line with the James-Lange theory of emotion (cf. Coleman & Snarey, 2011) and Damasio's (1994) somatic marker hypothesis. In addition, somatic and emotional experiences tend to interact, to the extent that both rely on active involvement of the anterior insula (Zaki et al., 2012). For example, pain perception is modulated by negative induced moods (Montoya et al., 2005; Villemure & Bushnell, 2002; Weisenberg et al., 1998). Stress and emotion also impact a variety of other somatic complaints, such as functional gastrointestinal disorders (Mayer et al., 2006).

To better study the interplay between emotions and interoception, the Affective Picture Paradigm was developed, a task assessing the influence of induced negative affect on symptom perception (Bogaerts et al., 2010). Viewing negative pictures led to more symptom reports, but only in high habitual symptom reporters (Bogaerts et al., 2010; Bogaerts et al., 2015). This effect was more pronounced when the pictures are perceived as arousing (Constantinou et al., 2013), although studies did not find differences in sympathetic measures of arousal, such as heart rate, breathing frequency, or skin conductance level. Analysis of brain activation patters during this paradigm showed that both nociceptive and somatosensory processes are involved in the symptoms elicited

by negative picture viewing (Bogaerts, Van Den Houte et al., 2023). As not many studies have been conducted using this paradigm, replication studies are needed before the Affective Picture Paradigm can be applied in a larger variety of settings.

To this end, we designed two studies. The goal of the first study was to find out which version of the Affective Picture Paradigm is the most effective for eliciting symptom reports – blocked vs. massed, with an explicit nocebo instruction vs. without. Our hypothesis was that people with higher baseline levels of persistent symptoms would report more symptoms during the negative pictures of the experimental task than people with lower baseline symptom levels. We also expected high symptom reporters to have less precise interoceptive abilities but show more somatosensory amplification (Köteles & Witthöft, 2017). The second study was a large-scale replication and extension of the first study. We expected to find the same effects regarding symptom perception, arousal, and valence. Additionally, the large sample size would allow us to calculate structural equation models to explore possible associations between the PHQ-15 bifactor model of symptom distress as well as the somatoform spectrum within the HiTOP framework. Lastly, we wanted to examine whether having a functional disorder diagnosis would impact self-report or experimental measures.

General methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow JARS (Kazak, 2018). Both studies mentioned below were part of a larger project preregistered at osf.io (ETUDE, 2022). Our study data and analysis code are also available in this repository.

Ethical approval was obtained from the ethical commission at the Psychological Institute of Mainz University (protocol 2021-JGU-psychEK-020). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956673. This article reflects only the author's view, the Agency is not responsible for any use that may be made of the information it contains. The current study is part of the innovative training network ETUDE (Encompassing Training in fUnctional Disorders across Europe; https://etude-itn.eu/), ultimately aiming to improve the understanding of mechanisms, diagnosis, treatment and stigmatisation of Functional Disorders (Rosmalen et al., 2021).

All questionnaires, unless explicitly stated otherwise, were embedded in SoSciSurvey (Leiner, 2019).

Study 1

Methods of Study 1

Participants

We recruited participants by sending an invitation with the study details to the university's psychology student mailing list, as well as to the researchers' friends and family. Eligibility criteria were being between 18 and 70 years of age and having access to a laptop or desktop PC. Initially, we aimed at a sample size of $n = 120$, as this would lead to roughly 30 people per condition. To ensure this number was met, we planned to overrecruit by 10%−20%. Participants were offered research credits and the chance to win one of five online shopping vouchers worth 20€.

However, due to a programming error compromising half the conditions, we restarted recruitment, aiming for a sample size of at least 60 persons ($n = 30$ per length-condition). We planned to stop the study after noticing more than $60 + 10\% = 66$ persons had completed the study. Again, we sent an advertisement for the study to the student mailing list but added a note to please not participate if they had already done the study earlier. Next to informing friends and family, we sent the study details to every psychology student association in Germany, to several press outlets (merkurist.de, a local newspaper in the Mainz region; Psychologie Heute, a nationwide layman psychology magazine), to a senior student mailing list, to all participants from former studies at our department, and we asked colleagues at different universities to share the study. We also shared the survey on reddit, twitter, Instagram and SurveyCircle, an online platform for sharing surveys.

Data collection took place from October 2021 to January 2022.

Materials

See Supplement 1 for a list of additional measures that were not primary outcomes of this study and were only used for correlational analyses.

Patient Health Questionnaire – 15 (PHQ-15). Kroenke et al. (2002) originally developed the PHQ-15 as a tool to assess the severity of somatic symptoms in the past 4 weeks. It is part of the PHQ-D (Spitzer et al., 1999). The PHQ-15 consists of 15 items (somatic complaints such as stomach pain, headache, fainting spells) on a scale of 0 (*not bothered at all*) to 2 (*bothered a lot*). After all items are summed up, a person who scores 5, 10, or 15 is considered to have low, medium, or high symptom severity, respectively. The PHQ-15 is reliable and well validated (Han et al., 2009; Kroenke et al., 2002). The German version of the PHQ-15 was established by Löwe et al. (2002).

Toronto Alexithymia Scale (TAS-20). Originally developed by Bagby, Parker, et al. (1994), the TAS-20 is a 20-item instrument used to assess deficiencies in recognising and describing emotions. Each item can be answered on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*). The instrument consists of 3 subscales: Difficulty Identifying Feelings, Difficulty Describing Feelings, and Externally Oriented Thinking. After multiple successful validation studies, the TAS-20 is considered a reliable and valid measure (Bagby, Parker, et al., 1994; Bagby, Taylor, et al., 1994). The German version was established by Bach et al. (1996).

Interoceptive Accuracy Scale (IAS). This scale was established by Murphy et al. (2019) to subjectively measure interoceptive accuracy. It comprises 21 items (e.g. item 9: "I can always accurately perceive when I am going to sneeze") on a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*). Higher scores attest to higher levels of interoceptive accuracy. Convergent and divergent validity have been demonstrated (Murphy et al., 2020). For the German translation, we used a not yet published version that we translated ourselves.

Somatosensory Amplification Scale (SAS). The somatosensory amplification scale measures to which extent somatic symptoms are experienced as bother- or worrisome. It was originally developed by Barsky et al. (1990). The SAS is a very economic instrument as it spans only 10 items to which respondents can answer on a 5-point scale (1 = *Not at all true*; 5 = *Extremely true*). It has shown good convergent and divergent validity (Barsky et al., 1990; Speckens et al., 1996). The German version is by Doering et al. (2015).

Affective Picture Paradigm. The Affective Picture Paradigm is an experimental task developed to assess the influence of negative affective states on symptom reporting (Bogaerts et al., 2010). There are many similar versions of the task (e.g. Bogaerts et al., 2015; Constantinou et al., 2013), but they all have in common that there are at least two types of picture series (negative affective pictures vs. one or more control pictures) followed each by some form of symptom assessment as well as an assessment of valence/arousal.

In our study, we used 48 pictures from the International Affective Picture System (IAPS, Lang et al., 2008). The pictures were presented for 7 s each. We used two groups of pictures: negative-valence pictures and neutral-valence pictures. In order to ensure comparability, we selected 8 pictures including animals, 8 pictures depicting humans, and 8 pictures showing objects per affect state. The following IAPS pictures were used for the negative-valence condition: 1114, 1200, 1302, 1820, 1932, 2120, 2683, 2692, 2800, 2811, 6020, 6212, 6370, 6550, 9001, 9140, 9181, 9230, 9410, 9470, 9561, 9600, 9800, 9911. For the neutral condition, we used pictures 1121, 1560, 1600, 1670, 1675, 1850, 1942, 1947, 2025, 2102, 2190, 2191, 2272, 2305, 2396, 2480, 5740, 7002, 7004, 7036, 7041, 7205, 7217, and 7546. These pictures were always followed by a Self-Assessment Manikin Rating (SAM, Bradley & Lang, 1994), which is a language-free measure of valence and arousal. The manikin images were accompanied by a 9-point scale. Note that the valence and arousal measures in this study therefore do not correspond to "objective" valence and arousal (such as population means reported in companion publications of the IAPS), but the subjective valence and arousal ratings of the participants in that moment.

This task was programmed in Inquisit Lab (Millisecond, 2018) and hosted by the Inquisit Lab webplayer.

Checklist Symptoms Daily life – state version (CSD). This questionnaire is embedded in the Affective picture paradigm. It assesses the presence and intensity of current somatic symptoms, such as nausea or headaches. The original English version is by Wientjes and Grossman (1994) and the German version was authored by Sauer and Witthöft (2017). The German version consists of 25 items that can be answered on a 5-point scale (*not at all* to *extremely*). With a high correlation with the PHQ-15 and a high internal

consistency, the CSD can be considered valid and consistent (Walentynowicz et al., 2018).

Design and procedure

Participants who clicked the link to the study were first directed to a landing page with participant information. They were then asked for informed consent (including secondary data use) and some demographic questions. After this, participants were prompted to fill in their email address to receive a pseudonymised code. Upon receiving the code, the participants were directed to the Inquisit Web landing page where they had to install the webplayer. This directly opened the experiment for them. Participants were then asked to fill in their pseudonymized code.

Participants were distributed to 1 of 8 groups according to a sequential algorithm in the backend of Inquisit Web. Depending on their condition, participants were either presented with neutral or negative pictures first, had an explicit nocebo instruction (or not) preceding negative pictures ("You will now see pictures with negative emotional contents. Earlier studies have shown that pictures with negative contents can lead to acute negative bodily symptoms. We therefore ask you to concentrate both on the pictures as well as on the presence of negative bodily sensations") and had either multiple short blocks (4 blocks per valence with 6 pictures each) or long blocks (1 block per valence with 24 pictures). In all the short blocks, 2 animal, 2 person, and 2 object pictures were shown, in a randomised order. Therefore, we had a 2*2*2 design (first valence * presence of explicit nocebo * block length). For an example of a short block condition without explicit nocebo instructions and with negative pictures first, see Figure 1. Unfortunately, a programming error initially led to only 12 pictures being shown in the negative valence part of the long conditions. As a result, we recruited more participants after fixing this mistake.

If participants were in the long block conditions, they needed to answer the full 25-item CSD after both blocks. If they were in one of the short block conditions, they had to respond to one quarter of the questions after each block ($6 + 6 + 6 + 7 = 25$). In total, both long-block and short-block participants answered the complete CSD 3 times (baseline, after negative, after neutral).

After they completed the experimental part, participants were instructed to go back to their browser tab. Here, the participants were asked to fill out all above-mentioned questionnaires. Lastly, they could select whether they wanted to receive participation credits or take part in a raffle to win shopping vouchers.

To ensure that participants did not simply skip the experimental part, we built in a 6-minute buffer on the page that redirected them to Inquisit web. Only after these 6 min would they be able to access the next questionnaires. We reasoned that 6 min was the shortest possible time in which the experimental task could be completed. Therefore, we hoped to minimise the motivation to simply click through the whole survey just to receive participation credits. Furthermore, we contacted people who had not completed the questionnaires or had issues with the experimental task and asked them to complete their participation (It is possible to do this in a pseudonymised fashion within SoSciSurvey).

Statistical analysis

The data was analysed using SPSS version 27 (IBM Corp, 2017) and R version 4.1.2 (R Core Team, 2013). After removing duplicate entries and recoding all

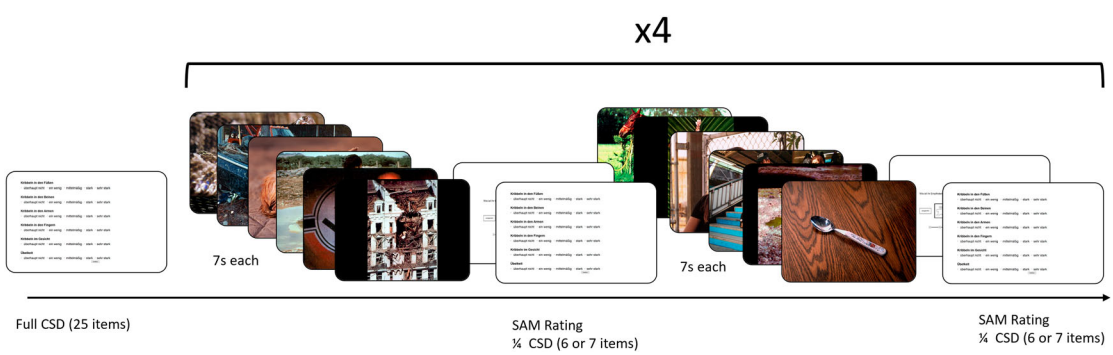


Figure 1. Illustration of one condition of the Affective Picture Paradigm.

inverted items, we first calculated reliability scores for all questionnaires using Cronbach's Alpha, as well as for the CSD-quarters in the negative and neutral trials in the short block conditions (a split-quarter reliability analysis, so to speak). The same was done for the SAM ratings across trial type.

We created a variable called CSD symptom-difference score (being the sum score of all CSD items after negative trials minus all CSD items after neutral trials) to collapse this information into one variable.

We conducted randomisation checks by conducting 3 pairwise *t*-tests (one per factor) and expected them to be non-significant. We additionally checked whether the version with the programming error was equal to the improved version by calculating *t*-tests over all relevant outcome measures in the long conditions (i.e. symptoms, arousal, and valence ratings after negative and after neutral trials = 6 *t*-tests). Another *t*-test was used to compare neutral-first vs. negative-first conditions with regard to their symptom scores. We further used *t*-tests to see whether the explicit nocebo conditions elicited more CSD symptoms than without explicit nocebo instructions. We also compared short vs. long block conditions with regard to symptom scores using a *t*-test. We tested the potential interaction effect between block length and presence of nocebo instructions using an ANOVA.

Furthermore, we conducted a correlational analysis. For a post-hoc comparison, we were interested in characteristics of high symptom-reporters vs. low symptom reporters (based on the PHQ-15 cut-off of 10). Here, too, we conducted *t*-tests across all scales.

Results of Study 1

Demographic analyses

All relevant demographic information can be seen in Supplement 2. In our first recruitment phase, $N = 136$ people completed the study. Of these, 6 persons were removed due to technical issues or missing experimental data. In the second recruitment phase, $N = 79$ people additionally participated, but due to incomplete data or technical issues, $n = 8$ persons were excluded.

Reliability analysis

All Cronbach's Alpha values can be found in Supplement 1.

Randomisation and comparability check

Regarding the start order, 101 people started in the negative-first condition and 100 in the neutral-first condition. There was no indication of unsuccessful randomisation when testing the baseline CSD, $t(199) = 0.605$, $M_{dif} = 0.413$, 95% $CI = [-0.93, 1.76]$, $p = .546$, $d = 0.085$. For the presence of explicit nocebo instruction ($n = 106$ in the condition with, $n = 95$ in the condition without), there was also no indication of unsuccessful randomisation, $t(199) = 0.561$, $M_{dif} = 0.39$, 95% $CI = [-0.98, 1.75]$, $p = .576$, $d = -0.112$. The same was the case for baseline variations in CSD in the short versus long block conditions ($n_{blocked} = 106$; $n_{massed} = 95$; $t(199) = -0.761$, $M_{dif} = -0.55$, 95% $CI = [-1.91, 0.81]$, $p = .448$, $d = -0.108$). As for the massed version with the programming error vs. the long version without the programming error, no differences were visible for symptoms after negative trials ($M_{dif} = -0.02$, 95% $CI[-0.16, 0.11]$, $t(93) = -0.352$, $p = .726$, $d = -0.074$), symptoms after neutral trials ($M_{dif} = -0.02$, 95% $CI = [-0.10, 0.07]$, $t(93) = -0.347$, $p = 0.729$, $d = -0.073$), valence after negative trials ($M_{dif} = -0.42$, 95% $CI = [-1.07, 0.23]$, $t(93) = -1.278$, $p = .204$, $d = -0.27$), valence after neutral trials ($M_{dif} = -0.37$, 95% $CI = [-0.95, 0.21]$, $t(93) = -1.278$, $p = 0.728$, $d = -0.27$), arousal after negative trials ($M_{dif} = 0.25$, 95% $CI = [-0.47, 0.96]$, $t(93) = 0.687$, $p = .494$, $d = .0145$), or after neutral trials ($M_{dif} = 0.47$, 95% $CI = [-0.28, 1.22]$, $t(93) = 1.348$, $p = .181$, $d = 0.285$).

Manipulation check

A paired sample *t*-test on CSD scores indicated that the manipulation was successful: persons reported more symptoms after viewing the negative compared to neutral pictures ($M_{diff} = 3.37$, 95% CI [2.69, 4.05], $t(199) = 9.777$, $p < .001$, $d = 0.690$). This was also the case when comparing these conditions for valence ($M_{diff} = -1.96$; 95% $CI = [-5.75, -4.31]$; $t(200) = -17.559$, $p < .001$, $d = -1.239$) and arousal ($M_{diff} = 1.32$ 95% $CI = [2.81, 4.07]$, $t(200) = 11.503$, $p < .001$, $d = 0.811$). We used a GLM ANOVA to further compare the CSD scores at baseline, after neutral pictures and after negative pictures. There were significant differences between all three symptom scores, with a significant reduction of symptoms compared to baseline after neutral pictures and a significant increase after negative pictures ($F(1.745, 348.912) = 56.981$, $p < .001$, $\eta_p^2 = 0.864$, Greenhouse-Geisser corrected after Mauchley-Test for

sphericity was significant). The differences between baseline and neutral were $CSD_{baseneutr} = -.925$, $p = .003$, 95% $CI = [-1.533, -.318]$. The difference between baseline and negative conditions was $CSD_{baseneg} = 1.913$, $p < .001$, 95% $CI = [1.174, 2.653]$. Lastly, for negative vs. neutral trials, the difference was $CSD_{negneutr} = -2.839$, $p < .001$, 95%$CI = [-3.330, -2.347]$.

Hypothesis check

We did not find differences in symptom ratings between persons in the explicit nocebo conditions and those in conditions without, $M_{dif} = 0.39$, 95% $CI = [-0.98, 1.75]$, $t(199) = 0.561$, $p = .576$, $d = 0.08$. We also did not find differences in the symptom rating differences for people in the short block conditions compared with people in the long block conditions, $M_{dif} = -0.55$, 95% $CI = [-1.92,0.81]$, $t(199) = -0.795$, $p = .428$, $d = -0.112$. Furthermore, there were no main effects nor interactions of these two factors ($F(4, 197) = 0.371$, $p = .774$,$\eta_p^2 = 0.006$). Lastly, there was no order effect (negative or neutral first) for symptom ratings ($M_{dif} = 0.62$, 95% $CI = [-0.74, 1.98]$, $t(199) = 0.903$,$p = .368$).

Exploratory analyses

Correlation analyses are available in Supplement 4.

We grouped participants according to their PHQ-15 scores. People below 10 were considered low habitual symptom reporters (LHSR, $N = 145$) and persons scoring 10 or more were considered high habitual symptom reporters (HHSR, $N = 56$). We found significant differences in the CSD symptom-difference score between LHSR and HHSR ($M_{dif} = -2.61$, 95% $CI = [-4.06, -0.55]$, $t(76.27) = -2.610$, $p = .011$, $d = -0.481$). In the CSD score after neutral pictures, there was a significant difference, indicating significantly larger symptom reports in the HHSR compared to the LHSR group ($M_{dif} = -5.083$, 95% $CI = [-6.82, -3.34]$, $t(71.414) = -5.821$, $p < .001$, $d = -1.126$). This was also the case for the CSD score after negative pictures, $M_{dif} = -7.39$,95% $CI = [-9.94, -4.83]$, $t(72.237) = -5.761$, $p < .001$, $d = -1.104$. These differences were already visible in the baseline CSD score, in which the HHSR group showed already significantly higher CSD scores compared to the LHSR group ($M_{dif} = -5.108$,95% $CI = [-6.77, -3.44]$, $t(70.184) = -6.118$, $p < .001$, $d = -1.200$). We did not find a significant group * picture condition interaction ($F(1.527,195.433) = 0.943$, $p = .391$, $\eta_p^2 = 0.007$).

There were no significant differences between HHSR and LHSR in terms of valence differences (negative minus neutral trials: $M_{dif} = -0.047$, 95% $CI = [-0.04, 0.54]$, $t(199) = 0.188$, $p = .851$, $d = 0.030$) or arousal differences ($M_{dif} = -0.26$, 95% $CI = [-0.83, 0.31]$, $t(82.097) = -0.915$; $p = .363$, $d = -0.161$). We did not find differences in IAS scores ($M_{dif} = -0.77$, 95% $CI = [-4.15, 2.61]$, $t(199) = -0.446$, $p = .656$, $d = -0.070$), but there were significant differences in SAS scores ($M_{dif} = -4.60$, 95% $CI = [-6.13, -3.07]$, $t(199) = -5.913$, $p < .001$, $d = -0.930$). We also found significant differences in the TAS subscale Difficulties Identifying Feelings ($M_{dif} = -3.31$, 95% $CI = [-4.94, -1.68]$, $t(83.962) = -4.035$; $p < .001$, $d = -0.698$).

Discussion of Study 1

This study provides initial evidence that the Affective Picture Paradigm is a valid and reliable tool to elicit somatic symptoms.

In this study, we tested different versions (i.e. blocked vs. massed) of the paradigm. We did not find any meaningful differences between the different versions. As for the presence of a nocebo, for reasons of parsimony we suggest not to include a statement meant to explicitly elicit symptoms, as it does not add anything to the already strong Picture Paradigm effect.

Correlations revealed that measures of affective disorders were stably correlated with symptom perception questionnaires, which fits previous findings (e.g. Van den Bergh et al., 2017).

We also looked at whether there were differences between people with a normal-to-mild range of symptoms vs. people with a clinically significant range of symptoms. This effect was medium-sized, with a Cohen's *d* of −0.481 in the CSDdifference variable. The effects were even larger when looking at the individual variables of CSD after negative pictures, at baseline, and after neutral pictures. However, the groups did not differ significantly in terms of valence and arousal, suggesting that group differences regarding symptom reports in the CSD are not explainable by differences in affective responses.

Study 2

Study 2 was part of a more extensive experimental setup examining interoceptive and emotion

regulation abilities in people with and without persistent somatic symptoms. The study was preregistered at osf.io (https://doi.org/10.17605/OSF.IO/43PR9). As the complete experiment consisted of multiple behavioural and self-report measures, not all of which are relevant to this paper, parts of the study outside of the present scope will be reported elsewhere.

Methods

Participants

To recruit participants from as diverse backgrounds and socioeconomic statuses as possible, we advertised on the student mailing, released a press statement, informed several press outlets, informed multiple social and charitable organisations (such as the Red Cross, the local homeless shelter), and displayed posters all around the city and university. We also used Twitter and Reddit to increase the visibility of our study. As this study took place on site, we asked our participants for proof of sufficient vaccination/recovery status regarding Covid-19. The only other exclusion criteria were being below 18 or over 65 years of age (in line with other projects from the ETUDE network, Rosmalen et al., 2021), and having epilepsy or Parkinson's Disease. We specified the latter because we were worried that people with photosensitive epilepsy would get a seizure, or that people with Parkinson's would have difficulties with reacting due to motor symptoms.

As this study was part of a large multimethod correlational study, we determined that we would ideally need 250 participants, as this is recommended when computing multiple correlations and structural equation models to achieve at stable estimates of correlation coefficients (Schönbrodt & Perugini, 2013). To ensure we reached this sample size, we slightly overrecruited and ended the data collection at 265 persons. The participants were paid 30€ or given research credits for their participation in the study.

Data collection took place from February to August 2022.

Materials

Here, we also used the PHQ-15, SAS, IAS, and TAS-20. For their descriptions, see above.

HiTOP somatoform phase 1 items (HiTOP-SF1). As explained in the introduction, the HiTOP model takes a dimensional approach towards psychopathology and aims to find transdiagnostic markers of mental health conditions (Kotov et al., 2017). The HiTOP-SF1 consists of 52 items answerable on a 4-point scale (Sellbom et al., 2021). The items represent five somatoform spectrum constructs (bodily distress symptoms, health anxiety, disease conviction, conversion, and somatic preoccupation). This questionnaire was developed in a large collaborative effort by the somatoform subgroup of the Measures Development Workgroup of HiTOP (Simms et al., 2022). The German version of the HiTOP-SF1 demonstrated adequate psychometric properties (Hartmann et al., 2022).

Design and procedure

Interested participants sent an email or scanned the QR code provided on the posters. If they had not provided a phone number, we then asked them to do so. We conducted a short telephone screening with all potential participants, asking them only three questions: their age, whether they had epilepsy, and whether they had Parkinson's. If the participants were eligible, we made individual appointments with each of them.

Two days before the lab appointment, we sent out reminder emails containing a battery of questionnaires (all the questionnaires listed above and a preliminary informed consent). Additionally, we asked about demographic details and current somatic, psychological, and functional conditions.

By completing the questionnaires, the participants triggered an automatic email sending a randomly generated, 8–12-digit alphanumeric code which functioned as the participant ID for the rest of the study. This would enable us to send invitations to certain subgroups (such as high habitual symptom reporters) for future studies without violating pseudonymity. Participants were further instructed not to drink alcoholic beverages or consume recreational drugs for the last 24 h before the lab appointment.

After informed consent (including secondary data use), participants underwent two heartbeat detection tasks (approx. 20 min in total) that will be described elsewhere. Then, the participants performed the Affective Picture Paradigm. Based on the results of Study 1, we chose to present short blocks without explicit nocebo instructions and with negative picture series first to all participants. The rest of the Affective Picture Paradigm remained the same as in Study 1. After this task, participants completed an

emotion regulation task (approx. 45 mins, reported elsewhere).

The tasks were completed on a 1280*1024px screen with a 60 Hz refresh rate.

Statistical analysis

The same analyses were conducted as in Study 1. Additionally, we conducted two SEM analyses on the relations between symptom groups of the PHQ-15 and the HiTOP, respectively, and the Affective Picture Paradigm. For the PHQ-15 model, we modelled symptom groups according to the findings on bifactor models by Witthöft et al. (2013, 2016, 2020). In these studies, the authors found that the PHQ-15 consists of a general factor of symptom perception that can be thought of as a negative-affective and motivational component of symptom distress motivational tendency to attend to symptoms, and secondly symptom clusters that represent the purely more specific sensory component aspects of the respective symptom types. Note that this affective-motivational component is a crucial part of symptom experience, encompassing the experienced discomfort and the drive for symptom relief. By disentangling this affective-motivational factor from the sensory factor, this research was able to show that people with somatic symptom disorders have an enhanced attention to symptoms (motivational g-factor) but reduced sensory detail (symptom clusters). We wanted to conduct these same analyses in this sample and extend them by relating these factors to the affective picture paradigm symptom reports. Therefore, we created a latent Affective Picture Paradigm variable that is created from all four APP-blocks. We used a WLSMV estimator with theta parametrization, as the PHQ-15 items are ordinal/categorical.

For the HiTOP model, we wanted to use a similar approach. Initially, we considered including all items to compute bifactor models as in the PHQ-15 models, but this would have led to too many model parameters and degrees of freedom for our PCs to reasonably handle. Additionally, as the HiTOP-SF has a much broader spectrum and more intricate (e.g. the bodily distress scale reflects symptoms and thus this subscale has the same objectives as the PHQ-15). A more parsimonious model was chosen, as the restrictions placed on classical bifactor models are not necessary in this case. Therefore, we used manifest subscale scores and used a maximum likelihood mean- and variance-adjusted estimator to estimate the latent HiTOP score. The *modindices* function revealed that allowing a covariance between the Health Anxiety and Disease Conviction subscales would drastically improve the fit. Due to empirical unidentification and very few degrees of freedom, we had to choose which subscale not to include when investigating the relations between HiTOP and Affective Picture Paradigm. We chose to leave out the correlation between conversion and the Affective Picture Paradigm, as already the authors of the HiTOP-SF reported difficulties with this subscale (Watson et al., 2022).

Lastly, we compared people with and without a functional disorder (FD) diagnosis on task performance measures and checked how this distinction compared to the PHQ-15 cutoff.

Results of Study 2

Demographic analyses

Different numbers of participants completed each part of the experimental set-up – with over 300 people screened, we managed to complete a lab session with $N = 265$ persons. Some persons did not complete the online questionnaires ($n = 11$). As a result, we have full data of $N = 254$ persons. Additionally, $n = 28$ persons filled out the online questionnaires but did not come to the lab session – usually due to illness and difficulty finding a new appointment. These 28 people were included in questionnaire reliability analyses, but not in any other statistic.

Demographic information is available in Supplements 3 and 4.

Reliability analyses

See Supplement 2.

Manipulation check

Participants reported significantly more symptoms in the CSD after negative ($M_{neg} = 34.40$, $SD_{neg} = 6.664$) than after neutral stimuli ($M_{neu} = 31.56$,95% *CI* [2.35, 3.33], $SD_{neu} = 5.063$; $t(253) = 11.371$, $d = .714$, $p < .001$). Valence was lower (i.e. more negative) in negative trials ($M_{diff} = -1.75$,95% *CI* [−1.92, − 1.58]$SD_{diff} = 1.370$; $t(253) = -20.388$, $d = 1.279$, $p < .001$) and arousal was higher ($M_{diff} = 1.68$, 95% *CI* [1.01, 1.32, 3.33] $SD_{diff} = 1.239$; $t(253) = 14.966$, $d = .939$, $p < .001$. A GLM ANOVA revealed the same pattern as above: neutral trials elicited even less symptoms than were present at baseline, while negative trials elicited

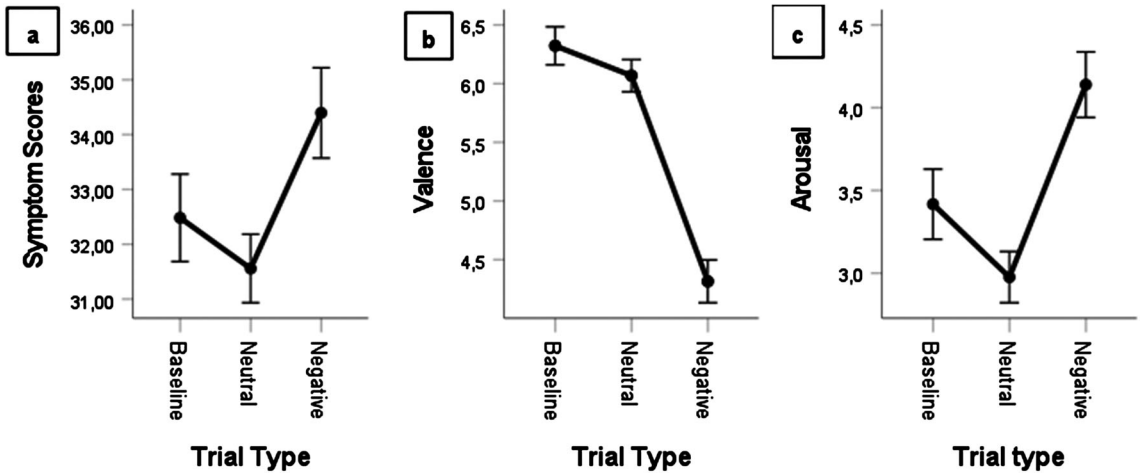


Figure 2. Visualisation of symptoms, valence, and arousal per trial type. Note. Error bars represent 95% confidence intervals.

the most symptoms ($F(1.652, 417.956) = 42.126$, $p < .001$, $\eta_p^2 = 0.143$, Greenhouse-Geisser corrected after Mauchley-Test for sphericity was significant). All contrasts were significant. The same patterns were evident when using valence as the outcome of the GLM ANOVA ($F(1.825, 461.840) = 311, 133$, $p < .001$, $\eta_p^2 = 0.552$, Greenhouse-Geisser corrected), or when using arousal as the outcome measure ($F(1.820, 460, 530) = 91, 960$, $p < .001$, $\eta_p^2 = 0.267$ Greenhouse-Geisser corrected). See Figure 2a–c for more information.

Exploratory analyses

Correlation analyses are available in Supplement 5.

$N = 199$ people were under the PHQ-15 cut-off of 10 (LHSR), and 55 scored 10 or higher (HHSR). HHSR showed a significantly larger difference in symptoms between negative and neutral trials than LHSR ($t(252) = -2.590$, 95% $CI = [-2.73, -0.37]$, $p = .01$, $d = -0.394$). These differences were also visible in baseline CSD ($t(59.035) = -4.385$, 95% $CI = [-8.93, -3.34]$, $p < .001$, $d = -1.030$), CSD after negative trials ($t(72.564) = -5.571$, 95% $CI = [-8.26, -3.91]$, $p < .001$, $d = -0.983$), and CSD after neutral trials ($t(67.235) = -4.998$, 95% $CI = [-6.34, -2.72]$, $p < .001$, $d = -0.961$). There were no significant group*picture interactions ($F(1.633, 411.469) = 2.853$, $p = .07$). Note that these statistics are adjusted for heteroscedasticity. There were no significant differences in valence difference ($t(252) = 1.312$, 95% $CI = [-0.14, 0.69]$, $p = .189$, $d = 0.201$) or arousal difference ($t(252) = -1.855$, 95% $CI = [-0.72, 0.02]$, $p = .065$, $d = -0.283$). HHSR showed significantly higher SAS scores than LHSR ($t(106.752) = -7.222$, 95%CI $[-7.78, -4.43]$, $p < .001$, $d = -0.964$) and scored significantly higher on the Difficulties Identifying Feelings subscale of the TAS-20 ($t(252) = -4.091$, 95% $CI = [-4.47, -1.57]$, $p < .001$, $d = -0.623$). There were no significant differences in IAS scores ($t(252) = 0.907$, 95% $CI = [-1.55, 4.20]$, $p = .365$, $d = 0.138$).

Regarding the Structural Equation Models, the baseline PHQ-15 model showed an adequate model fit ($\chi^2_{robust} = 84.365$, $df = 57$, $p = .011$; CFI = 0.996, RMSEA = 0.023, SRMR = 0.076) (see Figure 3). When including the Affective Picture Paradigm as a latent variable with its four CSD-difference measurements (one per block), the model fit was excellent ($\chi^2_{robust} = 124.254$, $df = 106$, $p = .109$, CFI = 1.0, RMSEA < 0.001, SRMR = 0.066) (Figure 4). Here, the significant correlations between the APP effect and the PHQ-15 g-factor ($r = .182$, $p = .036$) as well as the cardiorespiratory symptom factor ($r = .478$, $p < .001$) emerged as central findings.

The baseline HiTOP model also showed a good model fit ($\chi^2_{robust} = 6.954$, $df = 4$, $p = .138$; CFI = 0.982, RMSEA = 0.075, and SRMR = 0.030) (Figure 5). In the combined HiTOP-with-APP model, model fit was excellent ($\chi^2_{robust} = 21.965$, $df = 21$, $p = .402$, CFI = 0.993, RMSEA = 0.0026, SRMR = 0.033) (Figure 6). The correlation between the general HiTOP variable and the APP effect was significant ($r = .435$, $p = .018$). Exact coefficients are in Supplements 6–9.

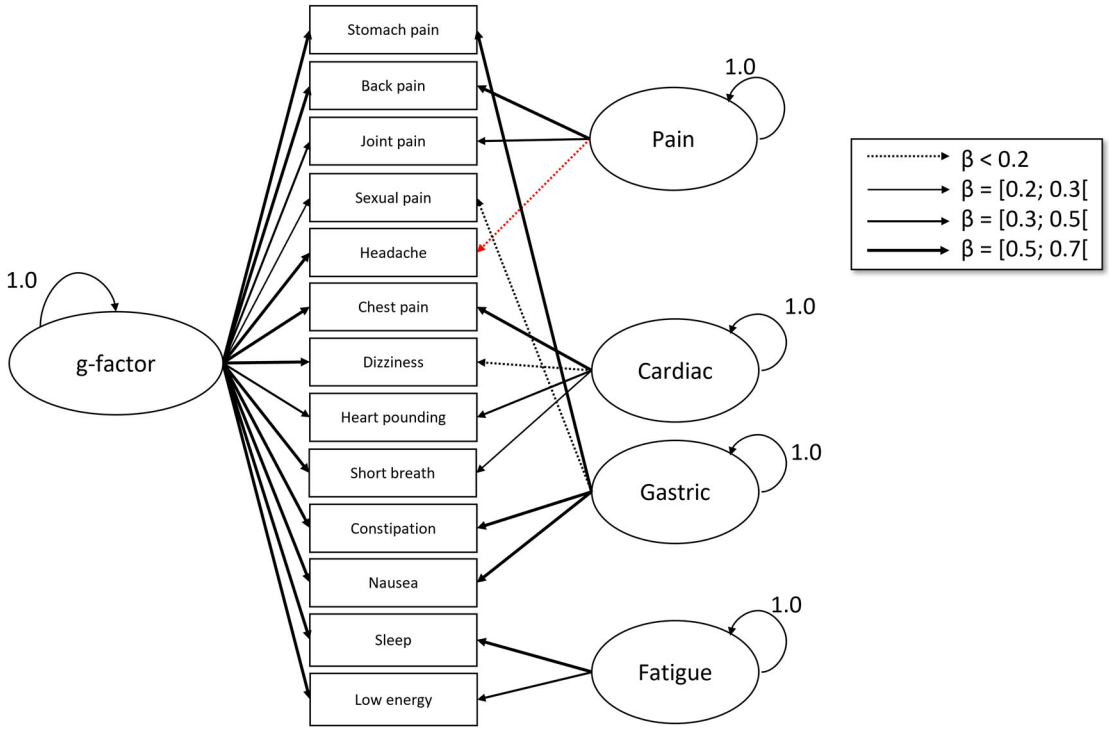


Figure 3. Baseline PHQ-15 bifactor model. Note. Single-headed arrows depict latent variable construction. Red arrows indicate negative relationships.

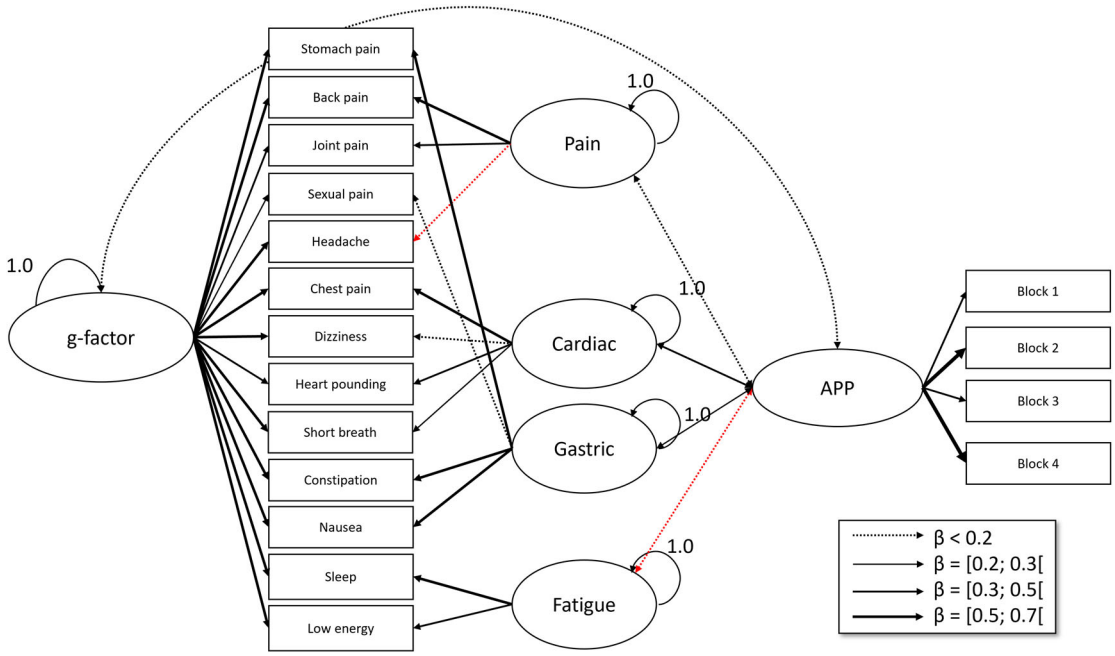


Figure 4. Structural equation model with PHQ-15 and Affective Picture Paradigm. Note. Double-headed arrows depict correlations, while single-headed arrows depict latent variable construction. Red arrows indicate negative relationships.

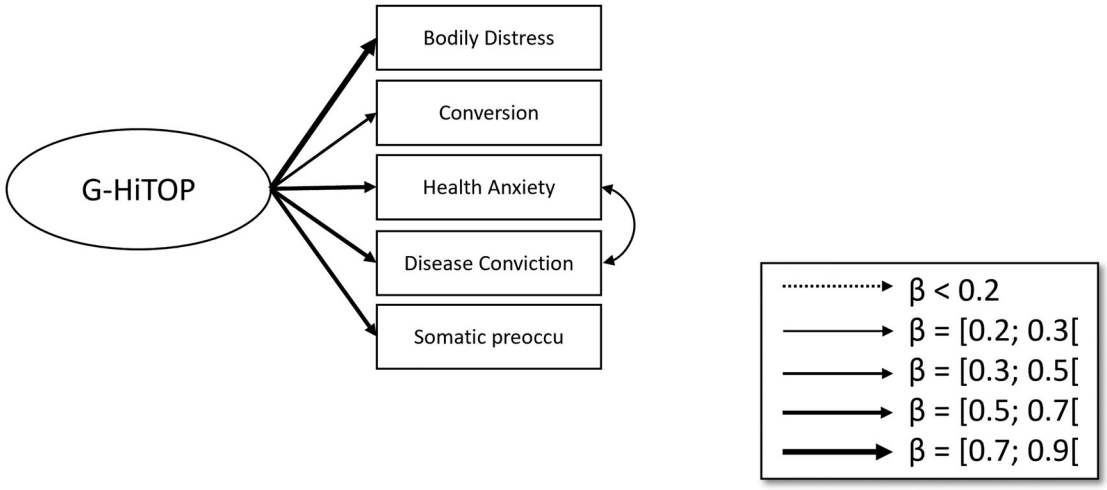


Figure 5. Baseline HiTOP model. Note. Double-headed arrows depict correlations, while single-headed arrows depict latent variable construction. Red arrows indicate negative relationships.

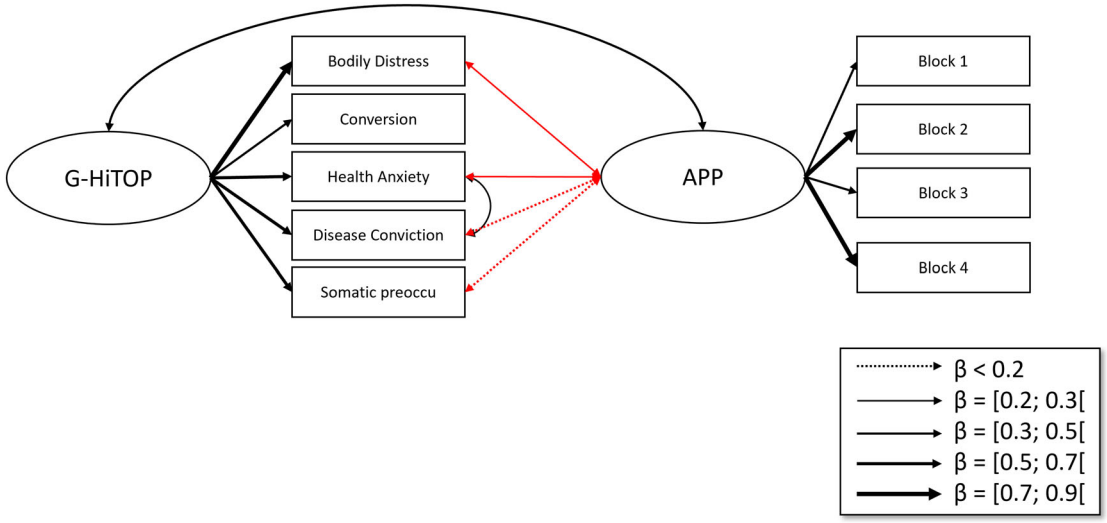


Figure 6. Structural equation model with HiTOP and Affective Picture Paradigm. Note. Double-headed arrows depict correlations, while single-headed arrows depict latent variable construction. Red arrows indicate negative relationships.

Functional disorder diagnosis

People with known or diagnosed functional disorders showed significantly stronger APP effects compared to participants without functional disorder diagnosis (CSD-difference variable) than people without ($M_{diff} = 1.87$, 95% $CI = [0.56, 3.18]$, $t(82.059) = 2.836$, $p = .009$). There were no significant differences in valence ($M_{diff} = -0.09$, 95% $CI = [-0.49, 0.30]$, $t(252) = -0.465$, $p = .642$) or arousal ratings ($M_{diff} = 0.256$, 95% $CI = [-0.08, 0.64]$, $t(252) = 1.548$, $p = .123$).

Comparing the FD diagnosis to the PHQ-15 cutoff showed $\chi^2 = 41,712$, $df = 1$, $p < .001$. Using the PHQ-15 as a shortcut for an underlying FD diagnosis correctly classified people in 79.13% of cases.

Discussion of Study 2

This study had two main aims: firstly, to replicate the finding of Study 1; and secondly, to improve our understanding of the mechanisms behind the Affective Picture Paradigm effect.

Regarding the first aim, largely the same patterns were found as in Study 1. The effect of picture viewing on symptom reporting in the CSD was present and strong: after negative pictures, symptoms and arousal were at their highest and valence at its lowest. The group differences were similar as in the earlier studies: People with PHQ-15 scores of over 10 showed significantly higher levels of symptom distress, but did not report feeling significantly more negative valence or higher levels of arousal than people with lower PHQ-15 scores. The correlations were also highly similar.

The Structural Equation Models showed that cardiorespiratory symptoms in the PHQ were most highly related to the Affective Picture Paradigm effect. In the HiTOP model, the general HiTOP factor showed a medium-sized correlation ($r = .435$) with the Affective Picture Paradigm effect. This suggests that the Affective Picture Paradigm effect is not specific to certain features or subfactors of somatoform pathology, but appears rather associated with the somatoform spectrum more broadly.

General discussion

The present studies shows that the Affective Picture Paradigm is a valid and economical tool for eliciting symptoms. We find evidence for the claim that negatively valent stimuli can induce the perception of bodily signals, which makes the Affective Picture Paradigm a valuable tool for experimentally studying (unexplained or persistent) somatic symptoms and its underlying mechanisms.

Our results also indicate that low and high symptom reporters perform differently on the task. This is important because one could assume that a symptom eliciting paradigm is not effective in people who, on average, already have high levels of symptoms, but the opposite is true – the Affective Picture Paradigm provokes even more symptoms in high habitual symptom reporters. This instrument might therefore be able to distinguish high and low habitual symptom reporters based on experimental performance, not just on plain self-reports.

We did not find any relationships or effects for the IAS questionnaire. Current research suggests that higher interoceptive accuracy coincides with or causes less symptom reporting (Schulz et al., 2020; Witthöft et al., 2020) and is generally protective of psychopathology (e.g. Bonaz et al., 2021). A possible explanation for this null finding is that interoception is difficult to measure and recent studies suggest that the different interoception questionnaires and experimental paradigms do not to measure the same underlying construct (Desmedt et al., 2022; Körmendi et al., 2022; Vig et al., 2022). A different approach to the measurement of interoception therefore may have yielded a different result.

Our work connects to earlier research indicating that trait negative affect is strongly involved in symptom perception (e.g. Van den Bergh et al., 2017). These correlations are consistent with clinically relevant comorbidities between affective disorders and functional disorders. Going one step further, we suggest that instead of two psychological conditions, they represent one underlying type of psychopathology that is characterised by these two features. The stable correlations of both affective and somatic measures with the HiTOP in Study 2 certainly point in this direction.

We did not find any significant differences in valence and arousal effects between high and low habitual symptom reporters, only symptom effects, which is consistent with earlier findings (Bogaerts et al., 2010; Bogaerts et al., 2015; Constantinou et al., 2013). Hence, high and low symptom reporters do not seem to perceive the stimuli differently. Instead, this supports the notion that the incoming (sensory) signals are very similar for low and high habitual symptom reporters – it therefore seems to be the prior (according to the predictive processing model) which causes the group differences in symptom reports. The same sensory input generates different posteriors – the experiences of symptoms. This is perfectly in line with the Predictive Coding Framework (Friston, 2010) and the Better Safe than Sorry Model (Van den Bergh et al., 2021). Continuing this line of thought, one may see a contradiction with the Somatic Marker Hypothesis (Damasio, 1994) or the earlier James-Lange-Theory (cf. Coleman & Snarey, 2011). These posit that the same number of somatic markers or visceral signals lead to similar affective processes in everyone. Arousal is one of the typical examples of somatic markers.

Interestingly, key mechanisms behind the APP effect seem to be cardiac and respiratory symptoms, according to our SEM models. Critchley and Garfinkel

(2017) have pointed out earlier that the cardiovascular channel is especially emotionally laden because of its particular innervation (e.g. baroreceptors, vagus nerve, etc.) and resulting need for interoceptive inhibition. This poor interoceptive inhibition would explain why people with more cardiorespiratory symptoms have a higher APP effect. It would be interesting to see in a direct comparison whether people who have FD but more gastric, fatigue-, or pain-dominated symptoms (e.g. fibromyalgia, irritable bowel syndrome), experience the APP-effect less strongly than people with cardiac functional disorders.

On a cognitive level, the SEMs with the HiTOP also show that the APP effect is most pronounced in people who have a wide variety of psychological symptoms associated with the somatoform spectrum. As a result, people who only fulfil one of the B criteria of the SSD diagnosis (e.g. only show health anxiety but not bodily distress, etc) will likely have a less pronounced APP effect – but still more than a healthy person. This once again highlights how useful the HiTOP might be in clinical practice.

A related finding was that high habitual symptom reporters have more difficulties identifying feelings (TAS-DIF subscale) than low habitual symptom reporters. Earlier studies have been able to show a moderating effect of negative affective pictures on symptom elicitation (Van Den Houte et al., 2017). Instead of arousal-perception (as stated by the Somatic Marker Hypothesis), affect (incorrect interpretation of bifactor/Better Safe Than Sorry Models), or interoception, *this* may be the key to the mechanism of symptom perception. As stated above, if interoception, arousal, and valence are similar across both low and high habitual symptom reporters, there must be a different top-down mechanism that is causing one group to experience symptoms while the other does not. Alexithymia might be the result of a stagnated error-reduction process as stated in the Predictive Coding Framework: although they are seeing the same pictures as low symptom reporters and attributing the same properties to them (i.e. same valence and arousal), as stated above, high symptom reporters might be unable to pinpoint that they are experiencing a diffuse emotional response to these pictures. (Negative) emotions usually have a physical component (e.g. feeling a lump in your throat when you are sad and want to cry, feeling nauseous when you feel shame or guilt). In high habitual symptom reporters, these physical components might be interpreted without considering the emotional context. Thus, these emotions are mistakenly labelled as somatic symptoms, as somatic priors are automatically advanced. A similar idea has been put forward by Jungilligens et al. (2022).

This explanation also aligns with findings from a recent brain imaging study (Bogaerts, Van Den Houte et al., 2023) which did not find differences between patients with FSS and healthy controls for a picture-induced neural emotion signature nor for a stimulus-intensity-independent pain signature. Instead, the difference between the two groups mediating the difference in symptom levels was a neural activation pattern involved in nociception (so-called neurological pain signature). This study showed that there are not necessarily affect-*induction* differences, but differences in neural *responses* towards the negative pictures. Apparently, they automatically trigger nociceptive priors. With our group comparisons, we arrive at a similar conclusion on a behavioural level.

While many different versions of this paradigm have been discussed in the literature (e.g. Bogaerts et al., 2010, 2022; Constantinou et al., 2013), we recommend using a short-blocked version without any explicit nocebo instructions and starting with a negative trial. This recommendation is based on parsimony (explicit nocebo was not significant) and experimental interest (e.g. short blocks allow the calculation of split-quarter reliability and a closer monitoring of the course of valence and arousal).

No study is without limitations – in our case, our samples consisted mostly of students or people with a highly educated background. Additionally, given the ethnic make-up of the Mainz region, most participants (likely) identified as white. However, both samples were large and quite heterogeneous when it came to important background characteristics such as age, gender, and habitual symptom reporting/functional disorder diagnosis. Additionally, we recruited in a large variety of settings (internet, homeless shelters, university) to enable as much sample diversity as possible. Thus, our findings have good generalizability. Another limitation is the programming error that occurred during Study 1. Usually, this would limit the internal validity, but with the far-from-significant comparability statistics and both study parts coming to highly similar results, we argue that this shows how strong the picture paradigm effect really is. Another strength is that our studies were conducted both online (Study 1) and in person (Study 2), which shows that the Affective Picture Paradigm is versatile, economic, and easy to use.

In conclusion, we showed that the Affective Picture Paradigm is a useful tool for the experimental provocation of somatic symptoms. Our findings shed light on the mechanisms behind symptom construction, which are in line with the Predictive Coding Framework. The differences between high and low habitual symptom reporters point to interesting avenues for future research on functional disorders.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by H2020 Marie Skłodowska-Curie Actions [grant number 956673].

ORCID

Tara M. Petzke http://orcid.org/0000-0003-2577-7695
Omer Van den Bergh http://orcid.org/0000-0001-6394-7363
Michael Witthöft http://orcid.org/0000-0003-1195-6326

References

Bach, M., Bach, D., de Zwaan, M., Serim, M., & Böhmer, F. (1996). Validierung der deutschen Version der 20-Item Toronto-Alexithymie-Skala bei Normalpersonen und psychiatrischen Patienten [Validation of the German version of the 20-item Toronto Alexithymia Scale in normal persons and psychiatric patients]. *Psychotherapie, Psychosomatik, medizinische Psychologie*, *46*(1), 23–28.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The twenty-item Toronto Alexithymia scale—II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, *38*(1), 33–40. https://doi.org/10.1016/0022-3999(94)90006-X

Barsky, A. J., Wyshak, G., & Klerman, G. L. (1990). The Somatosensory Amplification Scale and its relationship to hypochondriasis. *Journal of Psychiatric Research*, *24*(4), 323–334. https://doi.org/10.1016/0022-3956(90)90004-A

Bogaerts, K., Janssens, T., de van Peuter, S., Diest, I., & van den Bergh, O. (2010). Negative affective pictures can elicit physical symptoms in high habitual symptom reporters. *Psychology & Health*, *25*(6), 685–698. https://doi.org/10.1080/08870440902814639

Bogaerts, K., Rayen, L., Lavrysen, A., van Diest, I., Janssens, T., Schruers, K., & van den Bergh, O. (2015). Unraveling the relationship between trait negative affectivity and habitual symptom reporting. *PLoS One*, *10*(1), e0115748. https://doi.org/10.1371/journal.pone.0115748

Bogaerts, K., Van den Houte, M., Jongen, D., Ly, H. G., Coppens, E., Schruers, K., van Diest, I., Tack, J., van Wambeke, P., Petre, B., Kragel, P. A., Lindquist, M. A., Wager, T. D., van Oudenhove, L., & Van den Bergh, O. (2023). Brain mediators of negative affect-induced physical symptom reporting in patients with functional somatic syndromes. *Translational Psychiatry*, *13*(1), 285.

Bogaerts, K., Walentynowicz, M., Van den Houte, M., Constantinou, E., & Van den Bergh, O. (2022). The interoceptive sensitivity and attention questionnaire: Evaluating aspects of self-reported interoception in patients with persistent somatic symptoms, stress-related syndromes, and healthy controls. *Psychosomatic Medicine*, *84*(2), 251–260. https://doi.org/10.1097/PSY.0000000000001038

Bonaz, B., Lane, R. D., Oshinsky, M. L., Kenny, P. J., Sinha, R., Mayer, E. A., & Critchley, H. D. (2021). Diseases, disorders, and comorbidities of interoception. *Trends in Neurosciences*, *44*(1), 39–51. https://doi.org/10.1016/j.tins.2020.09.009

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Coleman, A. E., & Snarey, J. (2011). James-Lange theory of emotion. In S. Goldstein, & J. A. Naglieri (Eds.), *Encyclopedia of child behaviour and development* (Vol. 8, pp. 844–846). Springer US. https://doi.org/10.1007/978-0-387-79061-9_3146

Constantinou, E., Bogaerts, K., Van Diest, I., & Van den Bergh, O. (2013). Inducing symptoms in high symptom reporters via emotional pictures: The interactive effects of valence and arousal. *Journal of Psychosomatic Research*, *74*(3), 191–196. https://doi.org/10.1016/j.jpsychores.2012.12.015

Critchley, H. D., & Garfinkel, S. N. (2017). Interoception and emotion. *Current Opinion in Psychology*, *17*, 7–14. https://doi.org/10.1016/j.copsyc.2017.04.020

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 35–54). Springer Netherlands. https://doi.org/10.1007/978-94-017-9088-8_3

Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain* (1st ed.). Grossnet/Putnam.

Desmedt, O., Van den Houte, M., Walentynowicz, M., Dekeyser, S., Luminet, O., & Corneille, O. (2022). How does heartbeat counting task performance relate to theoretically-relevant mental health outcomes? A meta-analysis. *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.33271

Doering, B. K., Nestoriuc, Y., Barsky, A. J., Glaesmer, H., Brähler, E., & Rief, W. (2015). Is somatosensory amplification a risk factor for an increased report of side effects? Reference data from the German general population. *Journal of Psychosomatic Research*, *79*(6), 492–497. https://doi.org/10.1016/j.jpsychores.2015.10.010

Ehlers, A. (1993). Interoception and panic disorder. *Advances in Behaviour Research and Therapy*, *15*(1), 3–21. https://doi.org/10.1016/0146-6402(93)90001-I

ETUDE. (2022). *ESR1 Study 1*. https://doi.org/10.17605/OSF.IO/43PR9

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Garfinkel, S. N., Tiley, C., O'Keeffe, S., Harrison, N. A., Seth, A. K., & Critchley, H. D. (2016). Discrepancies between dimensions of interoception in autism: Implications for emotion and anxiety. *Biological Psychology*, *114*, 117–126. https://doi.org/10.1016/j.biopsycho.2015.12.003

Han, C., Pae, C.-U., Patkar, A. A., Masand, P. S., Woong Kim, K., Joe, S.-H., & Jung, I.-K. (2009). Psychometric properties of the Patient Health Questionnaire–15 (PHQ–15) for measuring the somatic symptoms of psychiatric outpatients. *Psychosomatics*, *50*(6), 580–585. https://doi.org/10.1016/S0033-3182(09)70859-X

Hartmann, J., Bräscher, A. K., Forbush, K. T., Sellbom, M., Watson, D., & Witthöft, M. (2022). The somatoform spectrum within the HiTOP system: A taxometric test of the latent structure. *Psychosomatic Medicine*, *84*(9), 1067–1076. https://doi.org/10.1097/PSY.0000000000001105

IBM SPSS Statistics for Windows (Version 25.0) [Computer software]. (2017). IBM Corp.

Inquisit Lab (Version 5.0.14.0) [Computer software]. (2018). Millisecond.

Jenkinson, P. M., Taylor, L., & Laws, K. R. (2018). Self-reported interoceptive deficits in eating disorders: A meta-analysis of studies using the eating disorder inventory. *Journal of Psychosomatic Research*, *110*, 38–45. https://doi.org/10.1016/j.jpsychores.2018.04.005

Jungilligens, J., Paredes-Echeverri, S., Popkirov, S., Barrett, L. F., & Perez, D. L. (2022). A new science of emotion: Implications for functional neurological disorder. *Brain: A Journal of Neurology*, *145*(8), 2648–2663. https://doi.org/10.1093/brain/awac204

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret, A. E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F. H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons, W. K., Stein, M. B., … Paulus, M. P. (2018). Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(6), 501–513. https://doi.org/10.1016/j.bpsc.2017.12.004

Körmendi, J., Ferentzi, E., Petzke, T. M., Gál, V., & Köteles, F. (2022). Do we need to accurately perceive our heartbeats? Cardioceptive accuracy and sensibility are independent from indicators of negative affectivity, body awareness, body image dissatisfaction, and alexithymia. *PLoS One*, *18*(7), e0287898. https://doi.org/10.1371/journal.pone.0287898

Köteles, F., & Witthöft, M. (2017). Somatosensory amplification – An old construct from a new perspective. *Journal of Psychosomatic Research*, *101*, 1–9. https://doi.org/10.1016/j.jpsychores.2017.07.011

Kotov, R., Jonas, K. G., Carpenter, W. T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K., Reininghaus, U., Slade, T., South, S. C., Sunderland, M., Waszczuk, M. A., Widiger, T. A., Wright, A. G. C., Zald, D. H., Krueger, R. F., Watson, D., & HiTOP Utility Workgroup. (2020). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry*, *19*(2), 151–172. https://doi.org/10.1002/wps.20730

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., … Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, *64*(2), 258–266. https://doi.org/10.1097/00006842-200203000-00008

Krueger, R. F., Hobbs, K. A., Conway, C. C., Dick, D. M., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Keyes, K. M., Latzman, R. D., Michelini, G., Patrick, C. J., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Tackett, J., Waldman, I., Waszczuk, M. A., … HiTOP Utility Workgroup. (2021). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry*, *20*(2), 171–193. https://doi.org/10.1002/wps.20844

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual: Technical Report A-8*.

Leiner, D. J. (2019). *SoSci Survey (Version 3.1.06)* [Computer software]. https://www.soscisurvey.de.

Löwe, B., Spitzer, R. L., Zipfel, S., & Herzog, W. (2002). *Gesundheitsfragebogen für Patienten (PHQ-D): Komplettversion und Kurzform* (2nd ed.). Pfizer.

Mayer, E. A., Naliboff, B. D., & Craig, A. D. B. (2006). Neuroimaging of the brain-gut axis: From basic understanding to treatment of functional GI disorders. *Gastroenterology*, *131*(6), 1925–1942. https://doi.org/10.1053/j.gastro.2006.10.026

Montoya, P., Sitges, C., García-Herrera, M., Izquierdo, R., Truyols, M., Blay, N., & Collado, D. (2005). Abnormal affective modulation of somatosensory brain processing among patients with fibromyalgia. *Psychosomatic Medicine*, *67*(6), 957–963. https://doi.org/10.1097/01.psy.0000188401.55394.18

Murphy, J., Brewer, R., Plans, D., Khalsa, S. S., Catmur, C., & Bird, G. (2020). Testing the independence of self-reported interoceptive accuracy and attention. *Quarterly Journal of Experimental Psychology*, *73*(1), 115–133. https://doi.org/10.1177/1747021819879826

Murphy, J., Catmur, C., & Bird, G. (2019). Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. *Psychonomic Bulletin & Review*, *26*(5), 1467–1471. https://doi.org/10.3758/s13423-019-01632-7

Rasmussen, E. B., & Rø, K. I. (2018). How general practitioners understand and handle medically unexplained symptoms: A focus group study. *BMC Family Practice*, *19*(1), 50. https://doi.org/10.1186/s12875-018-0745-2

R Core Team. (2013). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. http://www.R-project.org/

Rosmalen, J. G. M., Burton, C., Carson, A., Cosci, F., Frostholm, L., Lehnen, N., Olde Hartman, T. C., Rask, C. U., Rymaszewska, J., Stone, J., Tak, L. M., Witthöft, M., & Löwe, B. (2021). The European Training Network ETUDE (Encompassing Training in fUnctional Disorders across Europe): A new research and training program of the EURONET-SOMA network recruiting 15 early stage researchers. *Journal of Psychosomatic*

Research, *141*, 110345. https://doi.org/10.1016/j.jpsychores.2020.110345

Sauer, K. S., & Witthöft, M. (2017). Emotionserleben und somatische Beschwerden. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *46*(3), 147–156. https://doi.org/10.1026/1616-3443/a000427

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(*5*), 609–612. http://doi.org/10.1016/j.jrp.2013.05.009

Schulz, A., Rost, S., Flasinski, T., Dierolf, A. M., Lutz, A. P. C., Münch, E. E., Mertens, V.-C., Witthöft, M., & Vögele, C. (2020). Distinctive body perception mechanisms in high versus low symptom reporters: A neurophysiological model for medically-unexplained symptoms. *Journal of Psychosomatic Research*, *137*, 110223. https://doi.org/10.1016/j.jpsychores.2020.110223

Sellbom, M., Forbush, K. T., Gould, S. R., Markon, K. E., Watson, D., & Witthöft, M. (2021). Hitop assessment of the somatoform spectrum and eating disorders. *Assessment*, *10731911211020825*. https://doi.org/10.1177/10731911211020825

Simms, L. J., Wright, A. G. C., Cicero, D., Kotov, R., Mullins-Sweatt, S. N., Sellbom, M., Watson, D., Widiger, T. A., & Zimmermann, J. (2022). Development of measures for the Hierarchical Taxonomy of Psychopathology (HiTOP): A collaborative scale development project. *Assessment*, *29*(1), 3–16. https://doi.org/10.1177/10731911211015309

Sirri, L., Grandi, S., & Tossani, E. (2017). Medically unexplained symptoms and general practitioners: A comprehensive survey about their attitudes, experiences and management strategies. *Family Practice*, *34*(2), 201–205. https://doi.org/10.1093/fampra/cmw130

Speckens, A. E. M., Spinhoven, P., Sloekers, P. P. A., Bolk, J. H., & van Hemert, A. M. (1996). A validation study of the Whitely Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. *Journal of Psychosomatic Research*, *40*(1), 95–104. https://doi.org/10.1016/0022-3999(95)00561-7

Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*, *282*(18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737

Steinbrecher, N., Koerber, S., Frieser, D., & Hiller, W. (2011). The prevalence of medically unexplained symptoms in primary care. *Psychosomatics*, *52*(3), 263–271. https://doi.org/10.1016/j.psym.2011.01.007

Van den Bergh, O., Brosschot, J., Critchley, H., Thayer, J. F., & Ottaviani, C. (2021). Better safe than sorry: A common signature of general vulnerability for psychopathology. *Perspectives on Psychological Science*, *16*(2), 225–246. https://doi.org/10.1177/1745691620950690

Van den Bergh, O., Witthöft, M., Petersen, S., & Brown, R. J. (2017). Symptoms and the body: Taking the inferential leap. *Neuroscience & Biobehavioral Reviews*, *74*, 185–203. https://doi.org/10.1016/j.neubiorev.2017.01.015

Van Den Houte, M., Bogaerts, K., van Diest, I., de Bie, J., Persoons, P., van Oudenhove, L., & Van den Bergh, O. (2017). Inducing somatic symptoms in functional syndrome patients: Effects of manipulating state negative affect. *Psychosomatic Medicine*, *79*(9), 1000–1007. https://doi.org/10.1097/PSY.0000000000000527

Vig, L., Köteles, F., & Ferentzi, E. (2022). Questionnaires of interoception do not assess the same construct. *PLoS One*, *17*(8), e0273299. https://doi.org/10.1371/journal.pone.0273299

Villemure, C., & Bushnell, C. M. (2002). Cognitive modulation of pain: How do attention and emotion influence pain processing? *Pain*, *95*(3), 195–199. https://doi.org/10.1016/S0304-3959(02)00007-6

Walentynowicz, M., Witthöft, M., Raes, F., van Diest, I., & van den Bergh, O. (2018). Sensory and affective components of symptom perception. *Journal of Experimental Psychopathology*, *9*(2), jep.059716. https://doi.org/10.5127/jep.059716

Watson, D., Levin-Aspenson, H. F., Waszczuk, M. A., Conway, C. C., Dalgleish, T., Dretsch, M. N., Eaton, N. R., Forbes, M. K., Forbush, K. T., Hobbs, K. A., Michelini, G., Nelson, B. D., Sellbom, M., Slade, T., South, S. C., Sunderland, M., Waldman, I., Witthöft, M., Wright, A. G. C., … HiTOP Utility Workgroup. (2022). Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry*, *21*(1), 26–54. https://doi.org/10.1002/wps.20943

Weisenberg, M., Raz, T., & Hener, T. (1998). The influence of film-induced mood on pain perception. *Pain*, *76*(3), 365–375. https://doi.org/10.1016/S0304-3959(98)00069-4

Wientjes, C. J., & Grossman, P. (1994). Overreactivity of the psyche or the soma? Interindividual associations between psychosomatic symptoms, anxiety, heart rate, and end-tidal partial carbon dioxide pressure. *Psychosomatic Medicine*, *56*(6), 533–540. https://doi.org/10.1097/00006842-199411000-00009

Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., Olesen, J., Allgulander, C., Alonso, J., Faravelli, C., Fratiglioni, L., Jennum, P., Lieb, R., Maercker, A., van Os, J., Preisig, M., Salvador-Carulla, L., Simon, R., & Steinhausen, H.-C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*(9), 655–679. https://doi.org/10.1016/j.euroneuro.2011.07.018

Witthöft, M., Bräscher, A.-K., Jungmann, S. M., & Köteles, F. (2020). Somatic symptom perception and interoception. *Zeitschrift Für Psychologie*, *228*(2), 100–109. https://doi.org/10.1027/2151-2604/a000403

Witthöft, M., Fischer, S., Jasper, F., Rist, F., & Nater, U. M. (2016). Clarifying the latent structure and correlates of somatic symptom distress: A bifactor model approach. *Psychological Assessment*, *28*(1), 109–115. https://doi.org/10.1037/pas0000150

Witthöft, M., Hiller, W., Loch, N., & Jasper, F. (2013). The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *International Journal of Behavioral Medicine*, *20*(*2*), 172–183. http://doi.org/10.1007/s12529-012-9237-2

World Health Organisation. (2004). *International statistical classification of diseases and related health problems* (10th rev., 2nd ed.). Geneva: World Health Organization.

Zaki, J., Davis, J. I., & Ochsner, K. N. (2012). Overlapping activity in anterior insula during interoception and emotional experience. *NeuroImage*, *62*(1), 493–499. https://doi.org/10.1016/j.neuroimage.2012.05.012