OPEN

Creating Somatic Symptoms in the Lab: Stability and Predictive Validity of Symptoms in the Affect and Symptom Paradigm in a Sample From the General Population Over 18 Months

Tara M. Petzke, PhD, Ferenc Köteles, PhD, Omer Van den Bergh, PhD, and Michael Witthöft, PhD

Objective: The Affect and Symptom Paradigm (ASP) is an experimental setup that can reliably provoke somatic symptoms in a laboratory setting through the mere presentation of negative pictures. People with persistent somatic symptoms show characteristic effects in this paradigm, including elevated symptom reports and differential brain pattern activation, which suggests that the ASP might be a valuable diagnostic tool. In this cohort study, we tested the temporal stability and ability of the ASP to predict somatic symptom distress over a longer period of time.

Methods: We assessed $N = 91$ participants from the general population (69% female) over 2 time points, which were 18 months apart. Participants completed the ASP as well as a questionnaire on somatic symptom distress [Patient Health Questionnaire-15 (PHQ-15)]. Correlation analyses (Bayesian and frequentist) as well as cross-lagged-panel models were used to test the temporal stability and cross-lagged associations.

Results: In the cross-lagged-panel models, somatic symptom distress (PHQ-15) at T_1 significantly predicted ASP symptom provocation at T_2 ($\beta = 0.22$, $p = .029$) while controlling for the ASP at T_1. Moreover, ASP symptom provocation at T_1 significantly predicted cardiorespiratory symptoms (but not overall symptoms) in the PHQ-15 at T_2 ($\beta = 0.275$, $p = .019$). The autoregressive paths indicated moderate-to-high temporal stability (all $\beta > 0.27$, all $p < .050$).

Conclusions: The outcome of experimental somatic symptom provocation using the ASP appears stable over time and can significantly predict variability in the experience of cardiorespiratory symptoms (in the PHQ-15) 1.5 years later. Large cohort and intervention studies on chronic somatic symptoms and functional disorders may benefit from including experimental measures such as the ASP.

Key Words: persistent somatic symptoms, symptom perception, predictive processing, functional disorders

Abbreviations: (rm)ANOVA = (repeated measures) Analysis of variance, **ASP**, = Affect and Symptom Paradigm, **CLPM** = cross-lagged panel model, **CSD** = Checklist of Symptoms in Daily Life, **ETUDE** = Encompassing Training in fUnctional Disorders across Europe, **IAPS** = International Affective Picture System, **JARS** = Journal Article Reporting Standards, **PHQ** = Patient Health Questionnaire, **PSS** = persistent somatic symptoms, **SAGER** = Sex and Gender Equity in Research (Guidelines), **WEIRD** = westernized educated persons from industrialized rich democracies

(*Biopsychosoc Sci Med* 2025;87:453–461)

INTRODUCTION

Experiencing bodily symptoms over and over again can have debilitating effects on a person's well-being, and patients with persistent somatic symptoms (PSS) dare to dream about a future without these symptoms. PSS, defined as "subjectively distressing somatic complaints, irrespective of their etiology,"[1] affect up to two-thirds of the symptoms in primary care.[2] When PSS become debilitating and are accompanied by other symptoms, they can be classified as a somatic symptom disorder (according to DSM-5) or a functional disorder, which affect more than 8% of the general population[3] and have a large associated health care cost.[4] However, functional disorders and PSS are both usually seen through the lens of the specialty treating the affected organ[1]— for example, gastroenterologists treat and investigate irritable bowel syndrome, while neurologists treat and research functional movement disorders. While this specialization helps the research of the specific disorder at hand, this often leads to redundancies (ie, the

Received for publication January 16, 2025; accepted May 15, 2025.

From the Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, Johannes-Gutenberg-University Mainz (Petzke, Witthöft), Mainz, Germany; VIFASOM, Université Paris Cité (Petzke); Institut Psychiatrie et Neuroscience de Paris, INSERM U1266 (Petzke), Paris, France; Department of General Psychology and Methodology, Károli Gáspár University of the Reformed Church in Hungary (Köteles), Budapest, Hungary; and Health Psychology, University of Leuven (Van den Bergh), Leuven, Belgium.

Article Editor: Daryl O'Connor

Address correspondence to Tara M. Petzke, PhD, Psychological Institute, Wallstraße 3, 55122 Mainz, Germany. E-mail: tara.petzke@rub.de

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, journals.lww.com/bsam.

DOI: 10.1097/PSY.0000000000001413

same mechanisms are discovered in different disorders by different teams) and is to the disadvantage of people without a full clinical picture or whose symptoms fall "between the cracks." Newer research initiatives focus on interdisciplinary, transdiagnostic approaches to spearhead PSS research.[1,5]

One way to understand PSS is the active inference approach.[6–9] These models posit that body perception (and associated actions generated to contribute to perception) is shaped by hypotheses of the brain about the state of the body, as well as by physiological and context-related signals. If the hypothesis (prior belief or "prior") is highly precise while the somatic input has little precision (reliability), the eventual experience ("posterior") will be shaped more by the prior than by the actual somatic input. For example, negative affect as a trait can systematically bias priors about bodily dysfunction, which eventually contributes to the perception of low precise vague systemic responses as general malaise or fatigue. This process may further lead to a stagnated error reduction process, rendering PSS.[10]

An experimental paradigm that makes use of this finding is the Affect and Symptom Paradigm (ASP; formerly known as the "Affective Picture Paradigm"[11]. Because affective picture viewing using the IAPS picture set ("affective picture paradigm") has a long history in emotion research assessing verbal, psychophysiological, and brain responses to emotional pictures, whereas we modified affective picture viewing with IAPS pictures to specifically investigate symptom perception processes, we prefer to label this modified paradigm "the Affect & Symptom Paradigm" (ASP). In this experimental setup, negative (in comparison to neutral or positive) pictures are presented, followed by a symptom checklist. Non-consulting persons with high levels of chronic somatic symptom distress (sometimes termed "high habitual symptom reporters") typically report more elicited symptoms than "low habitual symptom reporters" in the ASP.[11–13] In line with the Predictive Processing Model, this is likely caused by a dominant symptom prior that reduces a person's reliance on sensory input. After all, if a person used their sensory input as the main source of information during this task, there would not be any marked symptom-level differences between negative and neutral trials. A recent paper showed that the effects of this paradigm are most strongly related to cardiorespiratory symptom perception.[14] A recent fMRI study with the ASP revealed that specific nociceptive (neurologic pain signature) and somatosensory (primary somatosensory cortex) brain activation patterns mediate symptom perception in functional disorder patients when viewing negative versus neutral pictures.[15] As these effects are especially strong in people with functional disorders and PSS, this suggests that the responses to the Affect and Symptoms Paradigm might constitute its own phenotype. It might even be a useful additional diagnostic aid, because this high susceptibility to symptom provocation by negative affect suggests a strong prior toward symptom perception and thus poses a heightened risk of developing PSS.

Few studies have studied experimentally induced symptoms over time. Diary studies have shown that prompting people to report their symptoms more frequently leads to an increase in symptoms over the course of a few weeks or months.[16,17] Even over the course of an 18-block ASP, asking about symptoms more frequently, immediately, and stably led to elevated symptom reports.[18] These studies show that the susceptibility to symptom provocation by negative affect stays more or less stable over the course of an experiment with consistent experimental manipulation, however long that experiment may be. In other words, there was no evidence of habituation or sensitization toward negative affective stimuli. The question remains whether a person's susceptibility to symptom provocation by negative affect changes over time, and which factors can influence this relation. Understanding this development could help clinicians intervene and potentially reduce this vulnerability, thus hopefully also alleviating people's symptom burden through adequate clinical interventions.

Thus, in the present study, we aimed to explore the interplay of symptoms elicited in the ASP and symptoms in everyday life [measured by the Patient Health Questionnaire 15 (PHQ-15)] as well as their temporal stability. As little is known about the temporal development of PSS, we chose an exploratory approach to best facilitate our research goals.

METHODS

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow JARS.[19] Our study is part of a larger project, and both the study design and analysis plan were preregistered at https://osf.io/sc57z/?view_only=3434eb2fcdfe4f37bbb436f02a7bfab2. Our study data and analysis code are also available in this repository (https://osf.io/gzp3u/?view_only=a6302d673a874636ba7c524237eca09e). Here, we look at the effects of the ASP without manipulation at T_1 and T_2, while in another article, we examined manipulations and covariates in a larger sample at T_2 only.[20] Another manuscript describing the full sample at T_1 is Petzke et al.[14]

Ethical approval was obtained from the ethical commission at the Psychological Institute of Mainz University (protocol 2023-JGU-psychEK). This project has received funding from the European Union's Horizon 2020 research and innovation program 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 stigmatization of functional disorders.[5]

Participants

For time point T_1, we recruited people via posters across the city, the university, and at humanitarian organizations, press releases, social media, and mailing lists within the university (see Petzke et al[14] for a full recap). Here, the goal was to recruit at least 250 persons, and we recruited $N = 265$, of which $N = 254$ could be included in the T_1 analysis. As stated in our preregistration, we attempted to re-recruit $N = 150$ persons for T_2 by notifying all former participants. However, our final sample size was $N = 91$. Despite our best attempts (eg, multiple emails and personally contacting people we knew had participated in T_1), we were not able to re-recruit more than these 91 persons from the T_1 assessment. Reasons for loss to follow-up included ineligibility ($n = 10$), incomplete pre-surveys and therefore not signing up for the study ($n = 8$), explicitly declining after completing the survey ($n = 2$), and no-shows ($n = 10$). However, we conducted a post hoc power analysis using G*Power version 3.1.9.7[21] and found that with $N = 91$, correlations at $|\rho| = 0.30$, and $\alpha = 0.05$, we had a sufficient statistical power of 84.3%.

At T_1, the exclusion criteria were having a diagnosis of epilepsy or Parkinson's disease and being younger than 18 years or older than 65 years. At T_2, the only exclusion criterion was being above 95 kg because of apparatus weight limitations.[20] Participants provided informed consent at both time points.

The total recruitment and data collection period for T_1 spanned from February 2022 to August 2022, and for T_2 from November 2023 to April 2024.

Materials

PHQ-15

The PHQ-15 is a 15-item instrument assessing symptom severity in the last 4 weeks.[22] Answers can be given on a 3-point scale (0 = *not bothered at all*; 2 = *bothered a lot*). This questionnaire has been validated in a variety of settings,[23–25] and has been translated into German by Löwe et al.[26] Reliability at T_1 was $\alpha = 0.78$ and at T_2 as well. The PHQ-15 was embedded in the SoSci Survey.[27]

Affect and Symptoms Paradigm (ASP)

Originally published by Bogaerts et al,[11] the ASP is a task in which participants are asked to rate the presence and intensity of bodily symptoms after watching negative or other pictures (from the International Affective Picture System, IAPS[28]). In the present study, participants viewed short blocks of 6 pictures for 7 seconds each (see Section 1 of the Supplemental Digital Content, http://links.lww.com/PSYMED/B103). There were 8 blocks: 4 negative blocks and 4 neutral blocks. Participants started by seeing a negative picture block first and then alternating between the 2 types of pictures, ending with a neutral picture trial. The task started with the 25-item Checklist of Symptoms in Daily Life—state version (CSD[29]), and a quarter of the CSD was presented after each block (6 questions after the first, second, and third block of both picture categories, 7 after the last block in both categories). In addition, participants indicated current valence and arousal at the beginning of the task on a 9-point Self-Assessment Manikin.[30] Participants were also asked to indicate their state of valence and arousal after every picture block. Participants needed ~15 minutes to complete the task, which was presented via Inquisit.[31] For more information, see Petzke et al.[14]

All stimulus materials and the order were kept the same at T_2, except that we shortened the symptom query list. This is described below. In this T_2 version, participants were given all 12 newly selected items after every block, rather than 6 or 7 different ones as in T_1.

Design and Procedure

Per STROBE guidelines, this is a cohort study.[32]

At T_1, participants completed a battery of questionnaires, including the PHQ-15 and some demographic questions, 2 days before their laboratory appointment. In the lab, participants were asked to complete the ASP and 3 unrelated other tasks.[33,34] Participants were compensated with 30€ or course credits for psychology students. At T_2, participants received a personalized email link to a survey with the above-mentioned questionnaire, as well as additional demographic and medical questions. This way, data from T_1 and T_2 could be combined in an anonymized fashion. Participants could then choose an appointment via a calendar tool. On the day of the appointment, participants completed 3 run-throughs of the ASP—before any manipulations, and 2 after 1 manipulation each. In this article, we only examined the ASP before any manipulation. Participants received 20€ or course credits.

Statistical Analysis

All analyses were conducted using SPSS 27,[35] JASP version 0.16.4,[36] and R 4.3.1.[37] Within R, we used packages foreign, psych, and lavaan.[38–40]

Based on the full $n = 254$ sample at T_1, we wanted to know which were the most provocative symptoms. We did this by using paired t tests per symptom between the negative and neutral trials in which the symptom was, and by examining effect sizes, consistency across studies we had performed, and content-wise overlap with the PHQ-15. This was done shortly after T_1 was finished.

After this, we examined all demographic markers and calculated questionnaire reliabilities using Cronbach's α in both samples. We calculated crosstabs to get a better overview of the health status of the participants. Then, we conducted a manipulation check using a repeated measures ANOVA to check whether baseline, negative, and neutral symptoms differed from each other. A Mauchly test for sphericity was conducted before every rmANOVA, and if it was significant, we used the Greenhouse-Geisser correction. Tests were 2-tailed, and significance was assumed at $\alpha = 0.05$. We calculated correlations using both a frequentist and a Bayesian approach. The primary outcome was the ASP effect, which is the difference between the sum of CSD symptoms after negative trials minus the sum of CSD symptoms after neutral trials.

Lastly, we used a data-driven approach to model cross-lagged effects of ASP and PHQ-15 scores. Cross-lagged panel models are a type of structural equation modeling that can be used to examine reciprocal longitudinal effects between (usually manifest) variables.[41,42] The 2 main types of paths are autoregressive paths (measuring how a construct develops over time based on its own previous measurements) and cross-lagged paths (measuring how earlier measurements of one construct influence later measurements of another construct). The time points are discrete; therefore, the term "panel." The results of cross-lagged panel models are usually reported in terms of β-values (standardized regression weights) as well as R^2 (explained variance). The β-values represent the effects of increasing "exposure" to the predictor by 1 unit.[43]

In our case, we modeled the longitudinal relationship over the 2 time points between the PHQ-15 and the ASP effect, which means that 4 paths were evaluated (2 autoregressive and 2 cross-lagged). We added the contemporaneous correlations between the 2 constructs at each time point. All 25 items were used for calculating the CSD/ASP-effect at T_1, while at T_2, we report the ASP-effect using the 12 items that we included.

Based on the first cross-lagged-panel model (CLPM), we decided to add another CLPM with the ASP effects and only cardiorespiratory PHQ-15 items as an exploratory analysis. We chose to examine these symptoms in particular as in an earlier study, cardiorespiratory PHQ-15 symptoms had the highest association with ASP effects.[14]

RESULTS

Demographic Information

At T_1, we recruited $n_1 = 254$ persons. Of those, $N = 91$ completed the T_2 assessment, and all data sets were included in the assessment. More women ($n = 63$, 69.2%) than men completed the study. The mean duration between the time points was $M = 569$ days (SD = 64), which is ~81 weeks. The flowchart of inclusion can be seen in Section 2 of the Supplemental Digital Content, http://links.lww.com/PSYMED/B103, and further demographic and medical information is shown in Table 1.

At T_1 and at T_2, $n = 18$ persons self-reported having a medical condition (excluding functional disorders). Crosstabs revealed that 6 persons no longer reported a medical condition, while 6 other persons contracted a medical condition ($\chi^2(1) = 31.09$, $p < .001$, $\varphi = 0.584$). For functional disorders, $n = 28$ people reported having one at T_1, but only $n = 18$ at T_2. Here, 14 people no longer reported an FD, while 5 other people newly reported functional disorders ($\chi^2(1) = 19.47$, $p < .001$, $\varphi = 0.463$). Lastly, psychological health also significantly improved over time ($\chi^2(1) = 43.748$, $p < .001$, $\varphi = 0.693$)—only 1 person newly reported a mental disorder, while 8 people no longer reported mental health conditions.

Regarding the PHQ-15, at T_1, participants had an average score of $M = 6.41$ (SD = 4.19), and at T_2, the average score was $M = 5.98$ (SD = 3.78). The cutoff for indication of persistent physical symptoms is at 5 points,[22] so most people in our sample were at or above this cutoff. In addition, note that this is slightly higher than for the general population (according to Hinz et al,[45] $M = 5.50$, SD = 3.93).

TABLE 1. Demographic Information

	M or *N* (SD or %)
Gender (male/female/other)	28/63/0
Age at T_1	28.95 (11.66)
Education	
No finished school form	1 (1.1%)
Low-level high school diploma	1 (1.1%)
Mid-level high school diploma	2 (2.2%)
High-level high school diploma	47 (51.6%)
Tertiary degree	40 (44.0%)
Occupation	
Student/in vocational training	62 (68.1%)
Employed	19 (20.9%)
Civil servants	1 (1.1%)
Self-employed	3 (3.3%)
Jobseeking/unemployed/rich	4 (4.4%)
Retired	2 (2.2%)
Self-reported psychological problems at T_2	13 (14.3%)
Depressive disorder	8 (8.8%)
Anxiety disorder	0
Other	7 (10.4%)
Self-reported structural physical conditions at T_2	18 (19.8%)
Neoplasms	1 (1.1%)
Endocrine, nutritional, and metabolic diseases	4 (4.4%)
Diseases of the eye and adnexa	3 (3.3%)
Diseases of the circulatory system	2 (2.2%)
Diseases of the respiratory system	2 (2.2%)
Diseases of the digestive system	1 (1.1%)
Diseases of the skin and subcutaneous tissue	2 (2.2%)
Diseases of the musculoskeletal system & connective tissue	5 (5.5%)
Congenital mal-/deformations, chromosomal abnormalities	1 (1.1%)
Allergies	1 (1.1%)
Self-reported functional conditions at T_2	18 (19.8%)
Irritable bowel syndrome	5 (5.5%)
Headaches	2 (2.2%)
Tinnitus	3 (3.3%)
Back pain (< 3 mo)	7 (7.7%)
Premenstrual syndrome	4 (4.4%)
Other	1 (1.1%)

Gender was assessed in accordance with the SAGER guidelines.[44] All participants' genders were the same as their sexes assigned at birth (no transgender individuals participated in this study).

Most Evocative Symptoms (in Negative vs. Neutral Trials)

As stated above, we analyzed which symptoms were most likely to be evoked by negative affect after T_1, which consisted of N = 254 participants, of whom all were invited to T_2. We also compared these *t* tests with an earlier study—experiment 1 in Petzke et al[14]—which consisted of $N = 201$ participants. The procedures in the 2 experiments reported in Petzke et al were similar: Experiment 1 was a pilot/feasibility trial where we wanted to find the most ideal setup of the ASP, while in experiment 2 (the T_1 of

this manuscript), we employed the ASP in a large, heterogeneous sample to examine associations with symptom burden and compute latent models. These 2 experiments were not entirely independent, as we recruited through the same means, but there were only $n = 8$ persons who participated in both trials.

Based on these results (see Section 3 of the Supplemental Digital Content, http://links.lww.com/PSYMED/B103), we decided that the 12 symptoms with which we would continue were *Shivering, Suffocating feeling, Need for air, Pressure on chest, Rapid heartrate, Feeling of heat, Feeling of head warmth, Hands tremble, Chest pain around heart region, Stiffness in fingers or arms, Pressure or knot in throat*, and *Faster/deeper breathing than normal*. These symptoms had good effect sizes across both studies and had some or high overlap with the PHQ-15 items. Note that especially the respiratory and cardiac symptoms were most subject to change across this and the previous study.

Manipulation Check

At T_1, there were significant differences between baseline, negative, and neutral trials regarding reported symptoms ($F(1.51, 135.91) = 26.30$, $p < .001$, $\eta^2_{part} = 0.226$). Negative pictures evoked most symptoms, followed by the baseline measurement, and lastly, the neutral picture condition. Similarly, at T_2, the baseline, negative picture, and neutral picture conditions differed significantly (regarding the reported symptoms) from each other ($F_{1.65, 148.04} = 28.79$, $p < .001$, $\eta^2_{part} = .242$). Here, the pattern was the same as at T_1, with the negative condition provoking the strongest symptoms.

Correlations

Correlations (frequentist and Bayesian) are presented in Section 4 of the Supplemental Digital Content, http://links.lww.com/PSYMED/B103. The pattern of correlations appears very similar between T_1 and T_2, with most associations being of medium size. Regarding stability of the ASP effect over time (with the shortened CSD at T_2), note that correlation with the longer CSD at T_1 was high ($r = 0.57$, $p < .001$) and that the 2 ASP versions had similar correlations with the other measures (eg, with the PHQ-15 at T_1: $r_{CSD_{T1}} = 0.24$, $p = .020$; $r_{CSD_{T2}} = 0.34$, $p < .001$).

Cross-lagged Panel Models

The first cross-lagged panel model, which included only PHQ-15 total scores and the ASP effect, can be seen in Figure 1. Here, the strongest predictors were autoregressions (for PHQ-15: $\beta_{PHQ} = 0.55$, $p < .001$, 95% CI = [0.37, 0.73]; for ASP: $\beta = 0.52$, $p < .001$, 95% CI = [0.39, 0.64]). The cross-lagged effect of the PHQ-15 total score at T_1 significantly predicted the ASP effect at T_2 ($\beta = 0.22$, $p = .029$, 95% CI = [0.07, 0.33]). Together with the cross-lagged regressions, $R^2 = 35.8\%$ of the variance in the PHQ-15 at T_2 could be explained, and $R^2 = 36.7\%$ of the variance of the ASP at T_2 could be explained. The second cross-lagged effect (of the ASP score at T_1 to the PHQ-15 total score at T_2) did not reach significance ($\beta = 0.14$, $p = .13$, 95% CI = [−0.060, 0.336]) but was descriptively in the expected (positive) direction. The size of the 2 cross-lagged coefficients did not differ significantly ($W = 0.0003$, $p = .99$).

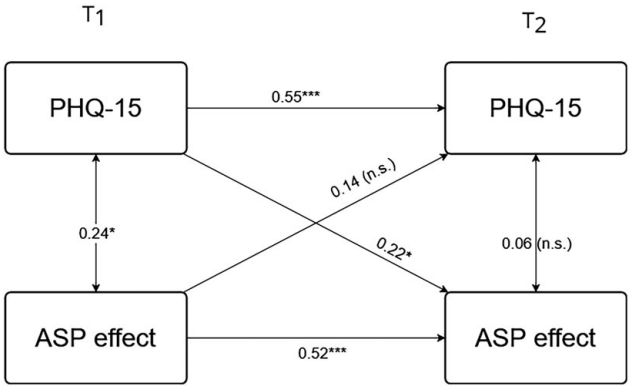

FIGURE 1. Cross-lagged panel model with PHQ-15 and ASP effect. Single-headed arrows depict regressions and have standardized regression weights, while double-headed arrows represent Pearson's correlations. *$p < .05$; **$p < .01$; ***$p < .001$. ASP = Affect and Symptom Paradigm; PHQ-15 = Patient Health Questionnaire 15 total score; CSD = difference score between negative and neutral trials on the Checklist of Symptoms in Daily Life; n.s. = not significant.

Because the ASP effect was found to be strongly associated with cardiorespiratory symptoms in a previous study, in our second CLPM, we only examined the cardiorespiratory symptoms in the PHQ-15, together with the ASP effect (Fig. 2). In this model, the autoregressive and cross-lagged components were more balanced—the ASP effect at T_1 significantly predicted the cardiorespiratory PHQ-15 scores at T_2 ($\beta = 0.28$, $p = .019$, 95% CI = [0.03, 0.52]) and vice versa ($\beta = 0.35$, $p = .050$, 95% CI = [0.12, 0.57]). Cardiorespiratory symptoms in the PHQ-15 (T_2)

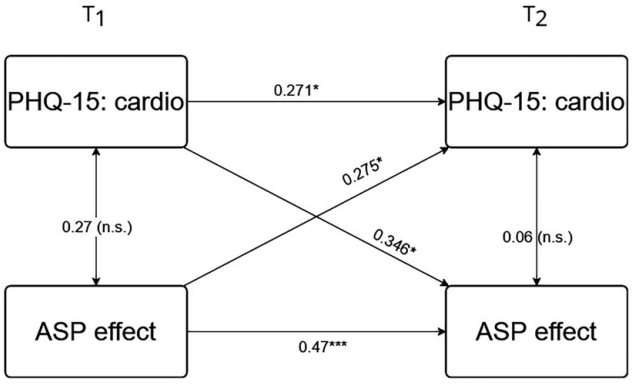

FIGURE 2. Cross-lagged panel model with cardiorespiratory PHQ-15 symptoms and ASP effect. Single-headed arrows depict regressions and have standardized regression weights, while double-headed arrows represent Pearson's correlations. *$p < .05$; **$p < .01$; ***$p < .001$. ASP = Affect and Symptom Paradigm; PHQ_cardio = Cardiac symptoms of the Patient Health Questionnaire 15; CSD = difference score between negative and neutral trials on the Checklist of Symptoms in Daily Life; n.s. = not significant.

were predicted with $R^2 = 19\%$ of explained variance, while $R^2 = 43.3\%$ of the ASP effect (T_2) could be explained.

DISCUSSION

The present study examined the temporal stability and predictive validity of the ASP across 2 time points that were 1.5 years apart. The 2 most important findings were the remarkably high degree of stability of the ASP effect (as an experimental measure of somatic symptom provocation in the lab) as well as its incremental predictive validity for the prediction of cardiorespiratory somatic symptom distress in the PHQ-15, even after statistically controlling for PHQ-15 baseline values at T_1.

Findings From the CLPMs

The cross-lagged panel models mostly strengthened the first impression obtained in the crosstabs—the autoregressions were medium high, but left some room for individual variation (ie, unexplained variation in the R^2), as also indicated in the turnover in people reporting a functional and/or structural medical condition. This fits with other findings that earlier symptoms predict later symptoms.[46,47]

The most important aspect of our CLPM analyses is that the cross-lagged paths are significant beyond within-variable temporal stabilities and beyond within-time point between-construct correlations. This means that the cross-lagged paths predict both the criterion variable and the changes in the criterion variable. Two interesting findings emerge from the cross-lagged regressions—first, earlier PHQ-15 total scores predict later symptom provocation. While earlier studies were able to show that the ASP has stronger effects in high symptom reporters (when habitual symptom reporting was assessed at the same measurement instance as the ASP[11–13]), this seems to be a time-stable finding. Earlier symptoms thus not only predict later symptoms, but also later susceptibility to symptom provocation (through negative affective stimuli).

Second, the outcome of the previous symptom provocation was able to predict later PHQ-15 cardiorespiratory symptoms. This fits with earlier findings that the ASP effect is particularly associated with cardiorespiratory symptoms in structural equation models.[14] An adjacent field to symptom perception is interoception, the processing of bodily stimuli.[48] Here, heartbeat tracking tasks and respiratory resistance tests are often used as a measure of how accurately a person is able to perceive their bodily states. There are still many open questions regarding how exactly cardiorespiratory perception should be measured and how it relates to symptom experience, but to many clinicians and researchers, it is clear that this modality plays a pivotal role in body perception.[49] Some examples include that heartbeat perception is linked to emotional experience with IAPS pictures,[50] and that (subjective) respiratory dysfunction is more pronounced in people with PSS.[51] Apparently, viewing emotional pictures generates cardiorespiratory symptom priors that subsequently determine how the brain disambiguates perception of the internal state.

This also would fit our suggestion from the introduction that scoring high on the ASP constitutes its own phenotype, or at the very least is an early warning sign for later PSS. In terms of the predictive processing model, reporting more symptoms after negative as opposed to neutral affective stimuli suggests that the symptom experience is primarily a result of a symptom prior, not "objective" somatic input.[15] Van den Bergh et al[10] illustrate this with the Better Safe Than Sorry Model: negative affect is a common underlying trait in people with PSS. This disposition can lead to bodily signals being classified as threatening and lead a person to attend more closely to their body, leading them to notice more threatening bodily signals—a vicious cycle. The ASP has the potential to help clinicians detect this disposition toward negative interpretation of bodily symptoms. Psychotherapists, psychiatrists, and other representatives of biopsychosocial medicine may find the idea of using a computerized test in a diagnostic setting a bit unusual, but in other areas of psychology and medicine, this is a standard practice (eg, in neuropsychology to detect the onset of neurodegenerative processes). We think that with its relatively short assessment time (< 20 min), the ASP could be a useful tool in clinical settings.

The Shortened CSD and Future Perspectives

In this study, we introduced a shortened 12-item version of the CSD,[29] which allows for a more economic assessment of symptoms in experimental setups such as the ASP. Especially in long experiments, participants often experience boredom when tasks are repetitive. This was the reason why we decided to split the 25-item inventory over the 4 blocks at T_1, but we decided we were not satisfied with this solution, as maybe certain stimuli provoked symptoms that were not asked for after that specific block. The 12-item version shows good convergent validity with the CSD at T_1, and similar correlation patterns with other symptom checklists, such as the PHQ-15. Therefore, we recommend using the shortened CSD in future ASP studies.

Given the predictive role of the ASP effect for future symptom experiences, the next step would therefore be to find interventions that can alleviate the susceptibility to or create a resilience against symptom provocation through negative affective stimuli in the ASP. One approach to this might be affective labeling, which has been shown to reduce symptom provocation in the ASP.[52] This is based on the finding that people with PSS score higher on alexithymia and often show higher levels of experiential avoidance.[53–55] People with a functional disorder have been shown to profit from Emotional Awareness and Expression therapy and experience symptom reduction.[56,57] Another approach has been physical activity to induce cardiorespiratory activation,[20] where ASP effects were reduced for the rest of the experiment in high symptom reporters. This led us to conclude that physical activity may reduce the activation of the symptom prior. We think it would be important to test more interventions within the ASP and translate successful findings into clinical exercises for affected people that they could easily do in their daily lives.

Constraints on Generality and Limitations

Our sample is a WEIRD sample (westernized, educated persons from industrialized, rich, democracies), which comes with many constraints on generalizability.[58] This was not entirely in our control, as it represents a completer sample (as opposed to an intention-to-test sample—and is therefore subject to a self-selection bias regarding dropout), and at T_1, we employed many different recruitment strategies to diversify our sample (such as contacting homeless shelters and charitable organizations). Furthermore, we did not assess the racial/ethnic identification or the cultural/geographic background characteristics of the participants, as these are socially unacceptable questions in Germany. The racial/ethnic makeup of our region is mostly white. However, we think that this sample is still valuable to study symptom perception—we have a quite high proportion of people with self-reported functional disorders, which was even higher than in the general population and thus allows us to study a broad range of symptom perception.[3] Similarly, in a normative study with over 9000 participants, the average value for the PHQ-15 was $M = 5.5$ (SD $= 3.93$), which is slightly lower than in our sample.[45] Lastly, Fischer et al[59] showed that student samples, which are similar to our sample—can be used to study functional symptoms.

A potential limitation of this study is that we were not able to recruit 150 participants at T_2 as planned. With 150 participants, we likely would have been able to run even more sophisticated models. However, as this research is exploratory, we were able to adaptively choose which models to run and had enough power to conduct these. For researchers planning a similar project, we recommend having a better plan regarding re-recruitment and preventing dropout at T_2 than we did. At T_1, we openly communicated with participants that they might be invited to a follow-up study, but we had not decided on the details and sample size yet. Potentially, having very concrete ideas in advance would make it more difficult to recruit participants at T1 (as, eg, students nearing graduation would likely reconsider) but reduce dropout rates.

Another limitation is that the symptom checklists at T_1 and T_2 were not exactly the same. As the T_2 version was also based on overlap with the PHQ-15, some bias may have been introduced in the cross-lagged effects. This issue, however, does not apply to autoregressive effects—first, while absolute scores of the ASP effect are not comparable between T_1 and T_2, regressions are based on the relative orders of predictor and criterion, not on absolute values, and thus allow appropriate interpretation. Second, if we had used a 25-item version at T_2 (whether once per block or split into quarters as at T_1), this would have led to higher shared variance between the 2 time points and thus our regressions would have had even higher parameter values. This further underlines the temporal stability of the ASP.

FUTURE DIRECTIONS AND CONCLUSIONS

For future research, it would be interesting to gather more longitudinal data with additional time points and then modeling more sophisticated latent models. One idea would be to use bifactor models of symptom experience, which have demonstrated that somatic symptom experience has affective-motivational and sensory components,[60–63] and see how the variables interact over time. This could be extended by using a random intercepts CLPM, which can model intraindividual fluctuations independently of the group mean and thus allows more precise conclusions on the level of the group and the individual.[64] However, this approach requires at least 3 waves of data. Taken together, these statistical approaches would allow us to gain a more fine-grained understanding of the mechanisms behind the ASP and the temporal predictors of symptom perception.

In conclusion, this study gives valuable insights into the temporal course of susceptibility to symptom provocation and hopefully provides food for thought regarding the development of PSS. Together with earlier studies, it strengthens the notion that the ASP could be a valuable diagnostic marker and aid in the early detection and prevention of PSS.

Source of Funding and Conflicts of Interest: This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 956673. The authors report no conflicts of interest.

Transparency and Openness Promotion: Our study design and analysis plan were preregistered at https://doi.org/10.17605/OSF.IO/sc57z. Our study data and analysis code are also available in this location.

REFERENCES

1. Löwe B, Andresen V, Van den Bergh O, Huber TB, dem Knesebeck OV, Lohse AW, et al. Persistent SOMAtic symptoms ACROSS diseases—from risk factors to modification: scientific framework and overarching protocol of the interdisciplinary SOMACROSS research unit (RU 5211). *BMJ Open*. 2022;12:e057596.
2. Steinbrecher N, Koerber S, Frieser D, Hiller W. The prevalence of medically unexplained symptoms in primary care. *Psychosomatics*. 2011;52:263–71.
3. Rometsch C, Mansueto G, Maas Genannt Bermpohl F, Martin A, Cosci F. Prevalence of functional disorders across Europe: a systematic review and meta-analysis. *Eur J Epidemiol*. 2024;39:571–86.
4. Konnopka A, Kaufmann C, König H-H, Heider D, Wild B, Szecsenyi J, et al. Association of costs with somatic symptom severity in patients with medically unexplained symptoms. *J Psychosom Res*. 2013;75:370–5.
5. Rosmalen J, Burton C, Carson A, Cosci F, Frostholm L, Lehnen N, et al. 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. *J Psychosom Res*. 2021;141:110345.
6. Bräscher A-K, Sütterlin S, Scheuren R, Van den Bergh O, Witthöft M. Somatic symptom perception from a predictive processing perspective: an empirical test using the thermal grill illusion. *Psychosom Med*. 2020;82:708–14.
7. Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci*. 2010;11:127–38.
8. Pezzulo G, Rigoli F, Friston K. Active Inference, homeostatic regulation and adaptive behavioural control. *Prog Neurobiol*. 2015;134:17–35.
9. Wolters C, Harzem J, Witthöft M, Gerlach AL, Pohl A. Somatosensory illusions elicited by sham electromagnetic field exposure: experimental evidence for a predictive processing account of somatic symptom perception. *Psychosom Med*. 2021;83:94–100.
10. Van den Bergh O, Brosschot J, Critchley HD, Thayer JF, Ottaviani C. Better safe than sorry: a common signature of general vulnerability for psychopathology. *Perspect Psychol Sci*. 2021;16:225–46.

11. Bogaerts K, Janssens T, Peuter S, De, Van Diest I, Van den Bergh O. Negative affective pictures can elicit physical symptoms in high habitual symptom reporters. *Psychology & Health*. 2010;25:685–98.
12. Bogaerts K, Rayen L, Lavrysen A, Van Diest I, Janssens T, Schruers K, et al. Unraveling the relationship between trait negative affectivity and habitual symptom reporting. *PLoS One*. 2015;10:e0115748.
13. Van Den Houte M, Bogaerts K, Van Diest I, Bie J de, Persoons P, van Oudenhove L, et al. Inducing somatic symptoms in functional syndrome patients: effects of manipulating state negative affect. *Psychosom Med*. 2017;79:1000–7.
14. Petzke TM, Weber K, Van den Bergh O, Witthöft M. Illustrating the pathway from affect to somatic symptom: the affective picture paradigm. *Cogn Emot*. 2024;38:1–17.
15. Bogaerts K, Van Den Houte M, Jongen D, Ly HG, Coppens E, Schruers K, et al. Brain mediators of negative affect-induced physical symptom reporting in patients with functional somatic syndromes. *Transl Psychiatry*. 2023;13:285.
16. MacKrill K, Groom KM, Petrie KJ. The effect of symptom-tracking apps on symptom reporting. *Br J Health Psychol*. 2020;25:1074–85.
17. Ferrari R, Russell AS. Effect of a symptom diary on symptom frequency and intensity in healthy subjects. *J Rheumatol*. 2010;37:2387–9.
18. Petzke TM, Elspaß L, Köteles F, Van den Bergh O, Witthöft M. And how did that make you feel?"—repeated symptom queries enhance symptom reports elicited by negative affect. *J Psychosom Res*. 2024;181:111634.
19. Kazak AE. Editorial: Journal article reporting standards. *Am Psychol*. 2018;73:1–2.
20. Petzke TM, Köteles F, Van den Bergh O, Witthöft M. Physical activity decreases somatic symptom distress in the affective picture paradigm. Submitted.
21. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41:1149–60.
22. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med*. 2002;64:258–66.
23. Han C, Pae C-U, Patkar AA, Masand PS, Woong Kim K, Joe S-H, et al. Psychometric properties of the Patient Health Questionnaire–15 (PHQ–15) for measuring the somatic symptoms of psychiatric outpatients. *Psychosomatics*.
24. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry*. 2010;32:345–59.
25. Stauder A, Witthöft M, Köteles F. Validation of the Hungarian PHQ-15. A latent variable approach. *Clin Neurosci/Ideggyogyaszati Szemle*. 2021;74:183–90.
26. Löwe B, Spitzer RL, Zipfel S, Herzog W. *Gesundheitsfragebogen für Patienten (PHQ-D): Komplettversion und Kurzform*, 2nd ed. Pfizer GmbH; 2002.
27. Leiner DJ. SoSci Survey (Version 3.5.02) [Computer software]. 2024. https://www.soscisurvey.de
28. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (iaps): affective ratings of pictures and instruction manual: Technical Report A-8. University of Florida, Gainesville, FL; 2008.
29. Wientjes CJ, Grossman P. Overreactivity of the psyche or the soma? Interindividual associations between psychosomatic symptoms, anxiety, heart rate, and end-tidal partial carbon dioxide pressure. *Psychosom Med*. 1994;56:533–40.
30. Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994;25:49–59.
31. Millisecond. *Inquisit Lab*. Millisecond; 2018.
32. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–7.
33. Petzke TM, Köteles F, Pohl A, Witthöft M. Somatic symptom distress is not related to cardioceptive accuracy. *J Psychosom Res*. 2024;181:111655.
34. Petzke TM, Witthöft M. The association of emotion regulation and somatic symptoms. *Psychosom Med*. 2024;86:561–8.
35. IBM Corp. *IBM SPSS Statistics for Windows*. IBM Corp; 2020.
36. JASP Team. JASP (Version 0.16.4) [Computer software]; 2022. Accessed November 24, 2022. https://jasp-stats.org/
37. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2013.
38. R Core Team. foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase'. 2022. Accessed July 13, 2022. https://CRAN.R-project.org/package = foreign
39. Revelle W. *psych: Procedures for Psychological, Psychometric, and Personality Research*. 2022. https://CRAN.R-project.org/package = psych
40. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Soft*. 2012;48:1–36.
41. Usami S. On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: interpretation of cross-lagged parameters and model choice. *Struct Equ Model Multidisci J*. 2021;28:331–44.
42. Kearney MW. Cross-lagged panel analysis. In: Allen M, ed. *The SAGE Encyclopedia of Communication Research Methods*. SAGE Publications, Inc; 2017.
43. Lüdtke O, Robitzsch A. *A Critique of the Random Intercept Cross-Lagged Panel Model*; 2021. doi:10.31234/osf.io/6f85c
44. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev*. 2016;1:2.
45. Hinz A, Ernst J, Glaesmer H, Brähler E, Rauscher FG, Petrowski K, et al. Frequency of somatic symptoms in the general population: normative values for the Patient Health Questionnaire-15 (PHQ-15). *J Psychosom Res*. 2017:9627–31.
46. Kitselaar WM, van der Vaart R, Perschl J, Numans ME, Evers AWM. Predictors of persistent somatic symptoms in the general population: a systematic review of cohort studies. *Psychosom Med*. 2023;85:71–8.
47. Creed F. The predictors of somatic symptoms in a population sample: the lifelines cohort study. *Psychosom Med*. 2022;84: 1056–66.
48. Khalsa SS, Adolphs R, Cameron OG, Critchley HD, Davenport PW, Feinstein JS, et al. Interoception and mental health: a roadmap. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2018;3:501–13.
49. Critchley HD, Garfinkel SN. Interoception and emotion. *Curr Opin Psychol*. 2017;17:7–14.
50. Parrinello N, Napieralski J, Gerlach AL, Pohl A. Embodied feelings —a meta-analysis on the relation of emotion intensity perception and interoceptive accuracy. *Physiol Behav*. 2022;254:113904.
51. Niwa S, Fila-Pawłowska K, Van den Bergh O, Rymaszewska J. Respiratory dysfunction in persistent somatic symptoms: a systematic review of observational studies. *J Psychosom Res*. 2024;181: 111607.
52. Constantinou E, Van Den Houte M, Bogaerts K, Van Diest I, Van den Bergh O. Can words heal? Using affect labeling to reduce the effects of unpleasant cues on symptom reporting. *Front Psychol*. 2014;5:807.
53. Gucht V de, Heiser W. Alexithymia and somatisation: quantitative review of the literature. *J Psychosom Res*. 2003;54:425–34.
54. Panayiotou G, Leonidou C, Constantinou E, Hart J, Rinehart KL, Sy JT, et al. Do alexithymic individuals avoid their feelings? Experiential avoidance mediates the association between alexithymia, psychosomatic, and depressive symptoms in a community and a clinical sample. *Compr Psychiatry*. 2015;56:206–16.
55. Torunsky NT, Knauz S, Vilares I, Marcoulides KM, Koutstaal W. What is the relationship between alexithymia and experiential avoidance? A latent analysis using three alexithymia questionnaires. *Pers Individ Dif*. 2023;214:112308.
56. Lumley MA, Schubiner H, Lockhart NA, Kidwell KM, Harte SE, Clauw DJ, et al. Emotional awareness and expression therapy, cognitive behavioral therapy, and education for fibromyalgia: a cluster-randomized controlled trial. *Pain*. 2017;158:2354–63.
57. Maroti D, Lumley MA, Schubiner H, Lilliengren P, Bileviciute-Ljungar I, Ljótsson B, et al. Internet-based emotional awareness and expression therapy for somatic symptom disorder: a randomized controlled trial. *J Psychosom Res*. 2022;163:111068.

58. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci*. 2010;33(2-3):61–83; discussion 83-135.
59. Fischer S, Gaab J, Ehlert U, Nater UM. Prevalence, overlap, and predictors of functional somatic syndromes in a student sample. *Int J Behav Med*. 2013;20:184–93.
60. Cano-García FJ, Muñoz-Navarro R, Sesé Abad A, Moretti LS, Medrano LA, Ruiz-Rodríguez P, et al. Latent structure and factor invariance of somatic symptoms in the patient health questionnaire (PHQ-15). *J Affect Disord*. 2020;261:21–9.
61. Witthöft M, Bräscher A-K, Jungmann SM, Köteles F. Somatic symptom perception and interoception. *Zeitschrift für Psychologie*. 2020;228:100–9.
62. Witthöft M, Hiller W, Loch N, Jasper F. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. *Int J Behav Med*. 2013;20:172–83.
63. Witthöft M, Fischer S, Jasper F, Rist F, Nater UM. Clarifying the latent structure and correlates of somatic symptom distress: a bifactor model approach. *Psychol Assess*. 2016;28:109–15.
64. Hamaker EL, Kuiper RM, Grasman RPPP. A critique of the cross-lagged panel model. *Psychol Methods*. 2015;20:102–16.

