Nocebo hyperalgesia induced by social observational learning

Elisabeth Vögtle *, Antonia Barke, Birgit Kröner-Herwig

Georg-Elias-Müller-Institute for Psychology, Department of Clinical Psychology and Psychotherapy, University of Göttingen, Göttingen, Germany

Keywords:
Nocebo response
Social observational learning
Pressure pain
Pain catastrophizing
Verbal suggestion
Nocebo hyperalgesia

1. Introduction

The expectation that a treatment may cause certain symptoms can lead to the occurrence or worsening of those symptoms. This phenomenon is called the nocebo effect [4]. Nocebo effects have been shown to occur in a number of contexts, such as chronic illness [24,46,47,49], motor performance in Parkinson patients [6], the effects of opioids and sedatives in postoperative pain [5], the subjective experience during invasive procedures [45], and the occurrence of mass psychogenic illness [25]. In an experimental setting, nocebo effects on pain were observed with various pain stimuli like ischemic pain [6], electric stimulation [10,11,44], mechanical stimuli, or the application of histamine [44]. No empirical study was found in which nocebo effects on pressure pain were investigated.

Little is known of the possible routes by which nocebo responses are acquired. They can be induced by conditioning or verbal suggestion alone [11,44].

Observing others might be one way in which pain-related beliefs and attitudes are acquired [19]. A mother’s modelling had an impact on her child’s pain behaviour [18]. In the same way, the pain expression of adults was influenced by the observation of models [12,13]. It is unknown whether a nocebo response can be induced in this way. One study showed that placebo analgesia could be induced by social observational learning [9] and that the effect was associated with empathy. Mass psychogenic symptoms can be caused by observing another person demonstrating symptoms [25]. No study was found in which socially induced nocebo effects on pain were investigated.

Studies concerning factors influencing the nocebo response are rare. Women demonstrated more nocebo nausea after a conditioning procedure than after verbal suggestion alone, whereas men showed stronger responses to the verbal suggestion than to the conditioning procedure [22]. Studies of placebo analgesia showed that it may be linked to factors like sex and stress [2], fear of pain [26,27] and pain catastrophizing [40]. So far, the results are inconsistent [2] and the exact impact of these factors is unknown [15].

The aim of our study was to investigate whether a nocebo response can be induced by verbal suggestion as well as by social observational learning. We expected a positive correlation between the strength of the nocebo response and fear of pain and...
pain catastrophizing. For participants in the social observational learning group, we also expected a positive correlation between the nocebo response and empathy.

2. Materials and methods

2.1. Participants

Participants were recruited via advertisements on university billboards. A total of 85 female right-handed students participated in the study. Four participants were excluded because of chronic pain, chronic bladder infection, lupus, and arthritis, respectively. One participant discontinued the experiment because she found the stimulation too painful. The remaining 80 participants (mean age 22.5 ± 4.4 years) were assigned randomly to the 3 experimental conditions. The groups did not differ with regard to age (Table 1).

2.2. Design

A 3 × 2 repeated-measures design with the between-factor “condition” (verbal suggestion, social observational learning, control) and within-factor “ointment application” (yes, no) was used. Before the pain induction, the participants were randomly assigned to one of the three conditions. All participants received pressure pain stimuli on the ring, middle, and index finger of each hand. The order of the stimulation, the ointment application, and the nocebo side was fully randomized. The dependent measure was subjective pain intensity.

2.2.1. Conditions

White odourless ointment served as nocebo. It was presented in a neutral container like those used by dispensing chemists. The ointment was hypoallergenic and contained no fragrances, preservatives, dyes, or potentially irritating agents, but only the following ingredients: aqua, caprylic/capric triglyceride, glycerin, pentylene glycol, cocos nucifera, hydrogenated lecinthin, Butyrospermum parkii, hydroxyethylcellulose, squalane, sodium carbomer, xanthan gum, carborner, and ceramide 3. In every condition the ointment was applied to one hand. The conditions differed regarding the information that possible ways of standardizing pressure pain application were being investigated. The ointment would have no effect on pain experience, but only on the moistness and elasticity of the skin. Participants in the verbal suggestion condition (VSC) read a nonverbal instruction on the experimental procedure was being tested. To this end they would be asked to watch a video in which the procedure was demonstrated. The participants in this group received no further written information. To focus their attention on the video they received 5 questions about the video prior to watching it. They were told that they would have to answer the questions after the video.

2.2.2. Pain ratings and nocebo response

Participants were asked to indicate pain intensity verbally during the application of a pressure pain stimulus every 20 seconds on an 11-point numerical rating scale with 0 indicating “no pain at all” and 10 indicating the “worst pain imaginable.” The investigators wrote down the answers.

The nocebo response was defined as the difference in pain ratings between the pressure applications with and without ointment. Higher values indicated a stronger nocebo response.

2.2.3. Pain induction

Pain was induced by a stationary pressure pain algometer, which delivers constant pressure over a fixed period of time. The main part of the algometer consists of a lever with a weight. At its end, a plunger with a surface area of 3 mm² is located. When a button is pushed, the lever with the plunger is lowered onto the finger. It rises automatically after a predetermined period of time. As a safety measure, the procedure can be stopped immediately by the pushing of a button.

A weight of 300 g set at 8 cm was used, resulting in a total pressure of 0.92 MPa, which was applied for 60 seconds. This stimulation intensity was chosen because it was below the pressure intensities used in other studies [16,17] but above the pressure pain threshold found in other studies [7,31]. We tested the chosen intensity in a pilot study in order to ascertain that the pain was tolerable for 60 seconds for all participants.

2.2.4. Video

The video was taken in the same room in which the experiment took place. It showed a seated female model and the hands of a female experimenter. The video was taken from an angle behind the model so that the model’s face was only partially visible (for a still picture, see Fig. 1). The video lasted 10.22 minutes and the sequence was as follows: the experimenter told the model that pressure pain stimuli would be applied and explained the rating scale. The pressure application started with the right hand without ointment. The model rated the pain intensity verbally on the 11-point scale (see above). The model’s ratings for the hand without ointment ranged from 2 to 3. After all three fingers of that hand had been stimulated, the ointment was applied to the model’s left hand and allowed to take effect for 60 seconds. Then the application of pressure pain resumed. The model’s ratings for the hand with ointment ranged from 5 to 7.

<table>
<thead>
<tr>
<th></th>
<th>Verbal suggestion (n = 27)</th>
<th>Social observation (n = 26)</th>
<th>Control (n = 27)</th>
<th>F(2,77)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD</td>
<td>M  SD</td>
<td>M  SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAI-S</td>
<td>36.6 7.0</td>
<td>36.2 8.0</td>
<td>36.3 4.9</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td>STAI-T</td>
<td>40.1 8.7</td>
<td>41.2 8.3</td>
<td>41.6 8.5</td>
<td>0.20</td>
<td>0.82</td>
</tr>
<tr>
<td>CES-D</td>
<td>10.1 5.6</td>
<td>11.9 6.3</td>
<td>11.1 5.6</td>
<td>0.65</td>
<td>0.52</td>
</tr>
<tr>
<td>PCS</td>
<td>22.4 7.8</td>
<td>22.2 7.6</td>
<td>25.0 10.7</td>
<td>0.85</td>
<td>0.43</td>
</tr>
<tr>
<td>PASS</td>
<td>65.3 19.5</td>
<td>73.2 24.9</td>
<td>76.1 28.9</td>
<td>0.51</td>
<td>0.60</td>
</tr>
<tr>
<td>SSAS</td>
<td>27.1 4.7</td>
<td>27.0 6.0</td>
<td>27.7 6.5</td>
<td>0.10</td>
<td>0.90</td>
</tr>
<tr>
<td>Stress</td>
<td>3.1 1.9</td>
<td>3.4 1.9</td>
<td>3.0 1.6</td>
<td>0.38</td>
<td>0.69</td>
</tr>
</tbody>
</table>

ANOVA, analysis of variance; STAI-S, State-Trait Anxiety Inventory, state version; STAI-T, State-Trait Anxiety Inventory, trait version; CES-D, Center for Epidemiological Studies Depression Scale; PCS, Pain Catastrophizing Scale; PASS, Pain Anxiety Symptoms Scale; SSAS, Somatoform Sensitivity Amplification Scale.
2.3. Psychometric instruments

State and trait anxiety were assessed with the trait (STAI-T) and state version (STAI-S) of the State-Trait Anxiety Inventory [23]. Each version consists of 20 self-descriptive statements concerning different anxiety symptoms. For the state version, participants have to indicate how they feel at the present moment on a scale ranging from 1 (“not at all”) to 5 (“very much”). For the trait version, participants have to indicate how they feel in general on a scale ranging from 1 (“almost never”) to 5 (“most of the time”). In a sample of female students, both versions demonstrated good internal consistency in different situations (Cronbach $\alpha$ between 0.90 and 0.93 for the trait version and between 0.93 and 0.96 for the state version) [23].

Depression was assessed with the German short version [20] of the Center for Epidemiological Studies Depression Scale (CES-D) [37]. It is a 15-item self-report scale designed to measure depressive symptomatology in clinical and nonclinical populations. Participants had to indicate how frequently they had experienced a symptom in the past week on a scale ranging from 0 (“rarely” [<1 day]) to 3 (“most of the time” [5–7 days]). Internal consistency in a female population was good (Cronbach $\alpha = 0.91$) [20].

Pain anxiety was assessed with the Pain Anxiety Symptoms Scale (PASS) [28]. It consists of 40 statements concerning pain-specific anxiety symptoms. The symptoms can be combined into the 4 subscales “Cognitive Anxiety,” “Escape and Avoidance,” “Fear,” and “Physiological Anxiety.” Participants have to indicate on a scale ranging from 0 (“never”) to 5 (“always”) how frequently they experience each symptom. In a community sample, the questionnaire showed good internal consistency (Cronbach $\alpha = 0.74$ to 0.89 for the subscales and 0.93 for the total score, only female subsample). Convergent and discriminant validity estimates were also good [32].

To assess empathy, the German version of the Interpersonal Reactivity Index (IRI) [14] (in German: Saarbrücker Persönlichkeitsfragebogen [SPF] [34]), was used. The questionnaire consists of 16 items, which are answered on a 5-point scale ranging from 1 (“never”) to 5 (“always”). The scale “Empathic Concern” assesses other-oriented feelings like compassion and sympathy for the person in trouble. “Perspective-taking” describes the ability to put oneself in the position of another real person. The “Fantasy” scale describes the tendency to put oneself in the position of characters in books and films. Negative self-oriented feelings of fear and discomfort at witnessing the negative experience of another person are measured by the scale “Personal distress.” The global score is calculated by summing up the scales “Empathic Concern,” “Perspective-taking,” and “Fantasy.” The questionnaire shows good internal consistency (Cronbach $\alpha$ between 0.66 and 0.74 for the subscales) and its validity has been tested by comparing it with other empathy questionnaires as well as self-evaluation [34].

Catastrophizing thoughts concerning pain were measured with the German version of the Pain Catastrophizing Scale (PCS) [29,42]. It consists of 13 items that are answered on a 5-point scale ranging from 0 (“not at all”) to 4 (“all the time”). The subscale “Rumination” describes the inability to stop thoughts concerning pain. The items of the subscale “Magnification” reflect the tendency to exaggerate the threat value of pain stimuli. The subscale “Helplessness” describes the feeling of inability to deal with the pain. The PCS total score is calculated by summing up all the items. The questionnaire shows good internal consistency (Cronbach $\alpha = 0.66$ to 0.87 for the subscales and 0.87 for total PCS) [39].

Somatosensory amplification was measured with the German version of the Somatosensory Amplification Scale (SSAS) [48]. Somatosensory amplification describes the tendency to be hyper-vigilant regarding bodily sensations and to appraise them as abnormal or symptomatic of disease. The questionnaire consists of 10 statements that describe uncomfortable bodily sensations (e.g., being too hot or too cold, hunger contractions). Participants have to indicate how characteristic of them the statements are on a scale ranging from 1 (“not at all”) to 5 (“very much”). In a sample of students, the SSAS demonstrated reasonable convergent but weak discriminant validity [1]. The internal consistency in a patient sample was good (Cronbach $\alpha = 0.82$) [3]. As a measure of subjective stress, participants indicated on an 11-point scale how relaxed and calm (0) or tensed and nervous (10) they felt.

As a manipulation check, participants were asked to write down which hypotheses they thought were tested by the study and to rate the credibility of the cover story on a scale ranging from 0 (“not credible at all”) to 10 (“completely credible”).

2.4. Procedure

The study was approved by the local ethics committee and was carried out in accordance with the Declaration of Helsinki [38]. Participants were informed about the study and gave their written consent. They were tested by 3 female students in their final year, who were trained and supervised by the first and second authors (E.V. and A.B.). All participants, the model in the video, and the experimenters were female in order to eliminate the possibly confounding influence of sex that might otherwise arise from differences in pain sensitivity between healthy men and women [35] or the influence of experimenter sex on pain perception and reporting [21,36]. Gender differences in nocebo response may exist, but so far have been investigated only in connection with nausea [22]. In order to focus on the conditions central to this study, we decided to keep these factors constant by confining ourselves to testing women.

Participants were randomly assigned to one of the three conditions. All participants filled in the STAI-S and the two stress ratings. Then they received the instruction according to the allocated condition. Participants in the OLC watched the video and answered the questions concerning the video. Afterwards, the pain application started. Every participant received one pain stimulus on the middle phalanx of the ring, middle, and index finger of both hands and was asked to indicate the pain intensity. After the pain application, the participants filled in the remaining questionnaires (CES-D, STAI-T, SPF, PCS, PASS, and SSAS). Finally, they answered the questions for the manipulation check and rated the credibility of the cover story. In order to be able to investigate the nocebo response, it was necessary to employ a deceptive information procedure in that the participants were led to expect that they would receive an ointment with an active ingredient, although this was untrue. At the end, they were fully debriefed and given the opportunity to ask questions. The participants received course credits or 10 Euros for their time.
2.5. Hypotheses

The following interaction hypothesis was tested: in the VSC and OLC, pain intensity with ointment is higher than in the CC. Furthermore, for the VSC and the OLC, a positive correlation between the size of the nocebo response and pain anxiety, pain catastrophizing, and somatosensory amplification was expected. Only in the OLC was a positive correlation between the nocebo response and empathy expected.

2.6. Data analysis

Differences regarding age, stress, state anxiety, trait anxiety, depression, pain catastrophizing, pain anxiety, somatosensory amplification, and the credibility ratings among the three conditions were compared with one-way analyses of variance (ANOVA).

In order to test the hypotheses, the mean pain intensity score for each hand was calculated and a $3 \times 2$ repeated-measures ANOVA with the between-subject factor condition (CC, VSC, OLC) and the within-subject factor application of ointment (yes, no) was computed. Least significant difference (LSD) post hoc tests were used for further analyses. As measures of effect sizes, partial $\eta^2$ and Cohen's $d$ were calculated.

The nocebo response for each individual participant was determined by the difference between the mean pain intensity ratings with and without ointment for that participant. Higher values indicated a stronger nocebo response. Pearson correlations were computed for the VSC and OLC combined, as well as for the OLC only, between the nocebo response and pain catastrophizing, pain anxiety, and somatosensory amplification. The Pearson correlation between nocebo response and empathy was computed only for the OLC. The level of significance was set at $P < 0.05$. All analyses were carried out with Statistica for Windows software, version 10 (Statsoft Inc., Tulsa, OK, USA).

3. Results

The groups did not differ in stress, state anxiety, trait anxiety, depression, pain catastrophizing, pain anxiety, and somatosensory amplification (Table 1).

The $3 \times 2$ repeated-measures ANOVA for pain intensity showed main effects for ointment $[F(1,77) = 5.34, P = 0.023, \eta^2 = 0.06]$ and condition $[F(2,77) = 3.66, P = 0.030, \eta^2 = 0.09]$, and an interaction condition $\times$ ointment $[F(2,77) = 3.86, P = 0.025, \eta^2 = 0.09]$. LSD post hoc tests identified no significant differences in pain intensity with and without ointment within the CC ($P = 0.522, d = 0.09$) and the VSC ($P = 0.841, d = -0.03$). In the OLC, pain intensity was higher with than without ointment ($P = 0.0007, d = 0.52$). Also, pain intensity was higher in the VSC than in the CC when the ointment was applied ($P = 0.030, d = 0.61$) as well as when it was not applied ($P = 0.010, d = 0.66$). With ointment, pain intensity was at the same level for the OLC and the VSC ($P = 0.341, d = 0.27$). Without ointment, pain intensity was higher in the VSC than in the OLC ($P = 0.006, d = 0.76$). There were no differences in pain intensity between the OLC and the CC (with ointment: $P = 0.226, d = -0.35$; without ointment: $P = 0.823, d = 0.06$) (Fig. 2, Table 2).

For the nocebo conditions combined, the correlation between nocebo response and PCS score was significant: the higher the tendency for pain catastrophizing, the stronger the nocebo response. For the subscales, only the correlation between “Helplessness” and the nocebo response was significant. No correlation between the nocebo response and pain anxiety or somatosensory amplification was observed (Fig. 3A–D, Table 3). For the OLC, we also calculated correlations between nocebo response and the subscales of the IRI; none of them were significant (Table 4).

The one-way ANOVA of the credibility ratings for the information concerning the ointment revealed differences among the three conditions $[F(2,77) = 11.93, P = 0.00003, \eta^2 = 0.24]$. LSD post hoc tests showed that the credibility ratings were higher in the OLC ($M = 6.7, SD = 2.5$) than in the CC ($M = 4.1, SD = 2.6, P = 0.0003, d = -0.102$) and the VSC ($M = 3.6, SD = 2.2, P = 0.00002, d = -1.29$). The credibility ratings of the CC and VSC did not differ ($P = 0.47, d = 0.20$). For the VSC, the nocebo response correlated with the credibility rating ($r = 0.42, P = 0.031$), but not for the other two conditions (CC: $r = -0.31, P = 0.115$; OLC: $r = -0.15, P = 0.469$).

4. Discussion

We investigated whether a nocebo response to pressure pain could be induced by social observational learning and verbal suggestion. A nocebo response was induced by social observational learning. As hypothesized, participants in the OLC rated the pain stimuli with ointment as more painful than those without after they had watched a video in which a model displayed greater pain after the application of an ointment. Contradictory to our hypothesis, the verbal suggestion that an ointment would intensify discomfort and pain [12,13]. A more recent study demonstrated...
that children who observed their mother exaggerating her display of pain and indicating a low pain threshold reported a lower pain threshold than children of the control group whose mothers had no specific instructions [18]. These studies did not address nocebo effects, but they did indicate that observational learning can influence pain perception. The present study showed that observational...
Learning can result in nocebo hyperalgesia. It cannot be ruled out, however, that participants in the OLC displayed overt responses consistent with their perception of the social demands in the situation, that is, that the reporting behaviour was influenced by the wish to conform to that of the model and the pain perception was unaltered. It seems unlikely, however, that this was the only mechanism at work, as it cannot explain the correlation between pain catastrophizing and the nocebo response (see below).

As positive correlations of placebo analgesia with empathy in an observational learning paradigm had been reported [9], we decided to assess empathy. With regard to the nocebo response, however, no significant correlation between empathy and nocebo hyperalgesia was observed. Possibly, empathy is not as important for acquiring nocebo hyperalgesia as it is for acquiring placebo analgesia. This finding would be in line with results showing differences in the way in which a nocebo response is acquired compared with a placebo response [11]. Other explanations for this result cannot, however, be excluded. Only female students of psychology were tested, who uniformly showed high empathy as evidenced by the low SD of this measure across all participants. According to the author of the German version of the empathy questionnaire [33], the SDs in a sample of German adults aged 22 to 26 years range from 3.01 to 3.42, whereas the SDs for the subscales in our sample were markedly lower (1.6 to 2.2). Therefore, the nonsignificant correlation could also be owed to the limited variability in our sample.

Furthermore, there might have been a floor effect. Possibly, not much empathy is required to learn from pain experiences, especially if they are explicitly voiced, as the model's pain ratings were. If this is the case, it is likely that all our participants possessed the minimum level of empathy required. Furthermore, empathy may be more important in situations in which nonverbal pain behaviour has to be interpreted (e.g., facial expressions). Before discounting the role of empathy in the acquisition of nocebo hyperalgesia by observational learning, further experiments should investigate samples with more variability in empathy and avoid verbal indicators of pain experience.

The ratings of the participants in the VSC with and without ointment were as high as the ratings in the OLC with ointment. Even without ointment, participants in the VSC perceived the pain as very intense. No nocebo effect was observed in the VSC. This result is in contrast with the reports by Colloca et al. [11] and van Laarhoven et al. [44], who reported that nocebo effects were induced by verbal suggestion. A possible explanation is that the participants in the VSC failed to believe the information that the ointment would increase the sensitivity of the skin, which is supported by the analysis of the information's credibility. Participants in the VSC rated the instruction as less credible than participants in the OLC. In addition, in the VSC, the credibility correlated with the nocebo response, indicating that the participants' nocebo response varied depending on the degree to which they believed that the ointment would enhance skin sensitivity. The low credibility of the cover story might thus explain why no nocebo response was observed in the VSC and limits the conclusions that might be drawn regarding verbal suggestion. By itself, however, it does not explain why, regardless of the ointment, the pain ratings were as high in this condition than in both other conditions.

The participants had been randomly assigned to the conditions. They did not differ with regard to age, stress, state anxiety, trait anxiety, depression, pain catastrophizing, pain anxiety, and somatosensory amplification. Thus, these parameters cannot have influenced the overall high pain ratings in the VSC. It seems more likely that these were the result of an unintended effect of our instruction. Possibly, uncertainty about the aim of the study and the suggestion of pain might have led to anxiety, which in turn has been linked to nocebo hyperalgesia [4,8].

Pain catastrophizing, which is defined as a tendency towards exaggerated negative thoughts concerning actual or anticipated pain, has emerged as one of the best psychological predictors of pain-related outcomes [41]. Catastrophizers reported significantly greater pain intensity, more pain-related thoughts, and more emotional distress than noncatastrophizers during a cold-pressor task [42]. They focused more on the pain sensation and also rated other people's pain as more intense when they watched videos of people taking part in a cold-pressor task [43]. A positive correlation between the nocebo response and pain catastrophizing in the nocebo conditions was found in our study, in particular with the subscale “Helplessness.” Helplessness is conceptualized as a secondary appraisal process and describes the tendency to negatively appraise a person's ability to deal with pain [41]. The specific impact of helplessness on pain perception has not so far received much attention [30]. Our results suggest that the more helpless a person feels, the more susceptible she is to a nocebo suggestion.

Our hypotheses concerning the relationship between nocebo response and trait anxiety, somatosensory amplification, and fear of pain were not confirmed. No study was found in which the association between one of the three variables and the nocebo response was investigated. Associations between the three variables and the nocebo reaction still might be found in a less homogeneous sample, however. Further research is needed.

Some limitations of our study have to be mentioned. Only female students of psychology were investigated. Moreover, experimental pain stimuli were applied with the familiar trade-off of experimental control at the sacrifice of some ecological validity. A real-life pain stimulus is more threatening and ambiguous than a stimulus applied in a structured experimental setting. In addition, naturalistic pain behaviour might be more credible to the observers and therefore, exert a stronger impact on their own subsequent pain experience.

To sum up, in our experiment, we showed for the first time that nocebo hyperalgesia can be induced by social observational learning and that catastrophizing, especially helplessness, enhances the nocebo response. Further research in this area should ideally combine the positive aspects of an experiment with those of more naturalistic settings and systematically vary samples and nocebo induction procedures.

Conflict of interest statement
The authors have no conflicts of interest to report.

Acknowledgement
We gratefully acknowledge the help of Verena Gross, Magdalena Nickl, and Karin Krischke, who helped with the data collection.

References