From whom do we learn pain? The influence of the demonstrator's pain assessment skills on nocebo hyperalgesia induced via observational learning

Bajcar Elżbieta Anita[1*], Kłosowska Joanna[1], Meeuwis Stefanie[1], Rubanets Daryna[1,2], Badzińska Julia[1,2], Bąbel Przemysław[1].

[1] Jagiellonian University, Institute of Psychology, Pain Research Group, Kraków, Poland

[2] Doctoral School in the Social Sciences, Jagiellonian University, Kraków, Poland

*Corresponding author: Elżbieta A. Bajcar, Jagiellonian University, Institute of Psychology, Pain Research Group, ul. Ingardena 6, 30-060, Kraków, Poland. Tel: +48126632431; Fax: +48126632415; E-mail: elzbieta.bajcar@uj.edu.pl

Abstract

The study aimed to determine if the pain assessment skills of the individual being observed influence observationally induced placebo effects in pain. Participants were randomly assigned to two experimental groups and one control group. In the experimental groups, participants observed the responses to the placebo of a model who had previously shown either high accuracy (high-accuracy group) or low accuracy (low-accuracy group) in differentiating the intensity of pain stimuli. The model exhibited less pain in response to pain stimuli when a placebo was administered, and more pain when no placebo was given. The control group did not observe anyone experiencing pain during the experiment. All participants then received a series of pain stimuli of the same intensity, half of which were administered with a placebo, and rated the intensity of pain. Pain-related expectations of the participants were also measured. Observational learning caused a form of nocebo hyperalgesia in the participants, leading to an increase in pain in a context that had previously accompanied the model's heightened pain sensations. This effect was only observed in participants from the high-accuracy group. Observational learning resulted in pain-related expectations, regardless of the observed model's accuracy. However, these expectations mediated participants' responses to pain only in the low-accuracy group. The findings suggest that the pain assessment skills of the observed model play a significant role in shaping observers' pain-related expectations and their responses to a placebo.

Introduction

Over the past decade, research has shown that social factors contribute to the formation of placebo effects in pain [2,3,37,42]. These results indicate that pain relief (i.e., placebo hypoalgesia and analgesia) or pain amplification (i.e., nocebo hyperalgesia) following placebo intervention might result from prior observation of an analogous response of another person to a placebo.

Observational learning (OBL) effects are partially based on the observed model's attributes. Previous studies have shown that certain characteristics of the model, such as their sex [50], self-confidence [9], and social status [8], can contribute to the magnitude of observationally induced placebo and nocebo effects in pain. However, OBL appears beneficial when the observed person is perceived as having the skills to perform a given task in a given situation effectively [32,49]. Observing such a skilled model can increase situation predictability and raise the observer's self-efficacy, thus promoting learning [40,47]. The current study investigated whether the model's perceived accuracy in differentiating between pain stimuli of different intensities can influence the observers' responses to placebo. The model perceived as having poor pain assessment skills (i.e., low-accuracy model) may generate less-certain pain-related expectations in the observer than the model with such skills (i.e., high-accuracy model), thus contributing to weaker placebo responses. Based on the aforementioned data, we hypothesized that observing a model would elicit placebo hypoalgesia in observers (H1) and that a high-accuracy model would elicit a greater placebo effect than a low-accuracy model (H2).

Although pain-related expectations have been widely assumed to be involved in placebo effects in pain induced by OBL [2,3,15], their mediating role in shaping these effects has been confirmed only relatively recently [35,43]. One of the aims of the current study was to investigate the role of expectations in shaping placebo responses induced by observing low- and high-accuracy models. We hypothesized that

expectations elicited by observing low- and high-accuracy models would mediate the placebo effect (H3).

Prior studies have suggested that individual characteristics of the observer can contribute to placebo effects [17,30]. Therefore, in the current study the observers' empathy, optimism, fear of pain, sensitivity to bodily signals, and reward responsiveness were measured to explore their associations with the magnitude of the learning effect.

Materials and methods

Study design

The study protocol was approved by the Research Ethics Committee (Institute of Psychology, Jagiellonian University, Kraków, Poland, ref. no. KE/24_2021) and was preregistered on the Open Science Framework (https://osf.io/4w8b5). A mixed between-within-subjects randomized study design was applied, in which participants were assigned in a 1:1:1 ratio to one of the experimental groups: 1) high-accuracy group, 2) low-accuracy group, or 3) control group.

The experiment consisted of four phases: calibration, baseline, observational learning (OBL), and post-OBL. The OBL phase was conducted only in the experimental groups. It consisted of two stages (stimuli discrimination task and placebo induction), during which participants were shown two videos presenting the same model. The first video aimed to demonstrate the model's ability to discriminate the intensity of two thermal stimuli: participants in the low-accuracy group observed an ineffective model, while participants in the high-accuracy group observed an effective model. The second video showed the model responding with lower pain to thermal stimuli administered with a placebo, and responding with higher pain to stimuli applied without a placebo (see '*Procedure*'). An overview of the study's design and procedures can be found in Figure 1.

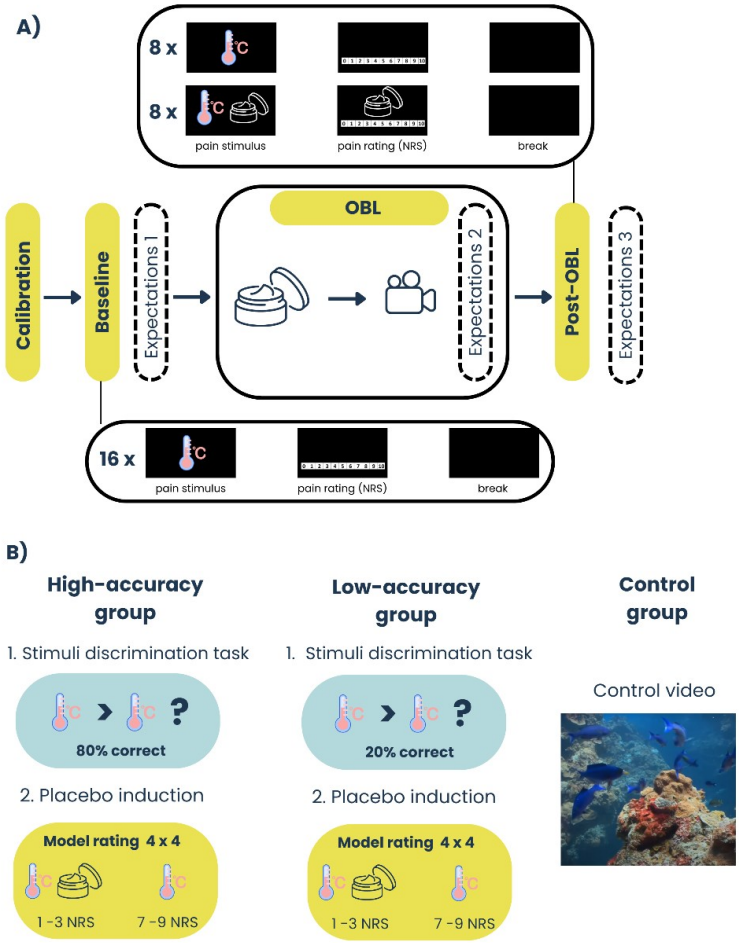

Figure 1. Study design. A) The experiment consisted of four phases: calibration, baseline, OBL, and post-OBL. During the baseline, participants received 16 pain stimuli of moderate intensity applied to both forearms. Then, a placebo ointment was applied to one of the forearms. During the OBL, participants observed either a video-recorded model (experimental groups) or a control video (control group). During post-OBL, 16 pain stimuli of moderate intensity were applied. Half of the stimuli were applied to the forearm covered with a placebo ointment. Pain expectations were measured after the baseline, after OBL, and after post-OBL. A single trial consisted of a pain stimulus accompanied by a thermometer icon presented on the screen, NRS for pain intensity rating, and a black screen before the next trial began. Notably, post-OBL trials included an additional ointment icon during thermal stimulus administration to the forearm covered with a placebo ointment and during the following NRS for pain intensity rating. **B)** In the experimental groups, OBL consisted of two videos: a stimuli discrimination task (accuracy manipulation) and a placebo hypoalgesia induction. In the first video, the model allegedly received two thermal stimuli and rated which one was more intense, providing

either 80% correct answers (high-accuracy group) or 20% correct answers (low-accuracy group). In the second video, the same model rated pain stimuli as weak with a placebo ointment and as strong without a placebo. Participants from the control group watched a control video of fish in an aquarium.
Note: OBL – observational learning; NRS – 11-point Numerical Rating Scale.

Participants

A power calculation for the mixed models was conducted using the superpower Shiny app [31]; https://arcstats.io/shiny/anova-power/) for the main effect of interest: a three-way interaction between experimental group (high-accuracy, low-accuracy, control), phase (baseline, post-OBL), and trials (placebo, control), with mean pain rating as a dependent variable. Power was calculated on 2000 simulated datasets for several possible values of the total sample size (i.e., n = 84, n = 96, n = 108, n = 120). The expected change in pain from the baseline to post-OBL in the experimental groups was determined based on prior findings regarding the effects of model characteristics (i.e., sex) on the efficacy of OBL in eliciting placebo effects [50]: $\Delta = 0.80$ in the high-accuracy model group, $\Delta = 0.50$ in the low-accuracy model group. Within-subject correlations between pain ratings and standard deviations were set at $r = 0.80$ and SD = 1.03, respectively. It was calculated that a sufficient power level ($1-\beta=0.83$) would be achieved with n = 96.

A total of 96 participants took part in the study, including 60 females (62.5%). Participants were recruited through advertisement websites and posters in various locations at Jagiellonian University. All participants were healthy volunteers aged 18–50 years of age, free of a history of chronic pain and any current or past neurological, psychiatric, or cardiovascular diseases. Volunteers were excluded if they had 1) participated in a pain study before, 2) studied psychology, 3) experienced pain on the day of participation, 4) consumed caffeine or nicotine at least 3 hours, alcohol at least 24 hours, or analgesics or psychoactive drugs 72 hours before their appointment. Each participant gave written informed consent and was compensated financially for their participation.

Apparatus and materials

Pain induction

Pain stimuli were applied to the volar surface of both forearms using two thermodes (active surface of 3x3 cm) connected to the Thermal Sensory Analyzer (TSA-II, Medoc Advanced Medical Systems, Ramat Yishai, Israel). In the course of the experiment, two thermal stimuli that elicited moderate pain were applied. The intensities of the two target thermal stimuli were calculated individually for each participant and separately for each forearm, based on the calibration procedure (see *Calibration*). All stimuli were initiated from 32.0 ^{0}C with a rate of temperature rise and fall of 10 °C/s. The target temperature was maintained for 6 seconds.

Placebo ointment

A neutral, colorless, and odorless ointment (Hasco-Lek, Wrocław, Poland) was used as a placebo. The application of the ointment was done on either the dominant or nondominant forearm, with the allocation counterbalanced among participants.

Measures

Pain intensity and pain expectation

Pain intensity and pain expectations were rated on an 11-point Numerical Rating Scale (NRS), ranging from “0” = “*no pain*” to “10” = “*the most intense pain that is tolerable*”. Pain ratings were collected after each thermal stimulus application, whereas expectations were rated after the baseline, after OBL, and after post-OBL.

Perceived accuracy

To assess the perceived accuracy of the model and participants’ predictions about their own accuracy, participants in the OBL groups were asked to rate on a 0–100 Visual Analog Scale 1) how accurately the model estimated the difference between the stimuli (0% ‘*not accurate at*

all'; 100% '*very accurate*'); 2) what percentage of the model's estimates were correct ('0% correct estimates'; '100% correct estimates'); 3) how accurate the participant would be when performing a similar task (0 '*not accurate at all*'; 100 '*very accurate*'). These questions were asked after the first video (presenting the stimuli discrimination task).

Physiological responses

The MP-160 hardware system and Acqknowledge v.5.0.2 software were used to collect electrocardiography (ECG) and electrodermal activity (EDA) data during the experiment (BIOPAC systems, Inc., Goleta, CA, USA). In order to prepare participants for ECG, their skin was abraded with Nuprep scrub (Weaver and Company, Aurora, CO, USA). Three disposable Ag/AgCl hydrogel (4% chloride salt) electrodes (Ø 11 mm) were connected to the leads: one each on the participant's left and right lower ribs, and one on the right side of the sternum. To measure the ECG signals, the ECG100D smart amplifier was used (gain x2000, high-pass filter of 1 Hz, low-pass filter of 35 Hz). For EDA, isotonic gel (0.5% chloride salt saline) was used to fill two disposable Ag/AgCi electrodes (gel cavity Ø 16 mm), which were placed on the medial phalanges of the index and middle finger of the non-dominant hand. To measure skin conductance level (SCL) and responses (SCR) to the heat stimuli, the EDA100D smart amplifier was used (low-pass filter of 3 Hz; IIR high-pass 0.5 Hz; cut-off, Q=0.707).

The PhysioData Toolbox 0.6.3 (Sjak-Shie, 2022) was used for visual inspection of the data and calculations of the data endpoints. Mean heart rate (HR; in bpm) and skin conductance level (SCL; in µS) were calculated for the baseline and post-OBL. In addition, the average skin conductance response amplitude (SCR; in µS) to the administration of the pain stimuli in the baseline and post-OBL was calculated. For SCR, a latency window of 1 to 5 seconds after the thermal stimulus onset was used.

Psychological traits

Empathy was measured by means of 1) the Interpersonal Reactivity Index, IRI [20]; Polish adaptation [29], which consists of 22 items divided into three subscales – Empathic Concern, Perspective Taking, and Personal Distress (without Fantasy Scale from the original IRI); 2) 22-item version of the Empathy Quotient, EQ short [56]; Polish adaptation [25].

Sensitivity to rewards and punishment was measured by a short form of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire [16]; Polish adaptation [57], which consists of 24 items divided into two subscales: Sensitivity to Punishment and Sensitivity to Reward.

Optimism was assessed by the 10-item Life Orientation Test-Revised, LOT-R [44]; Polish adaptation [26].

Fear of pain was measured by the 30-item Fear of Pain Questionnaire-III, FPQ-III [34]; Polish adaptation [33] with the following subscales: Fear of Severe Pain, Fear of Minor Pain and Fear of Medical Pain.

Attention to bodily signals was measured on the 16-item Body Attention, Ignorance, and Awareness Scale, BAIAS [7]; Polish translation was made for research purposes by the Pain Research Group), which includes the Body Attention, Body Ignorance, and Body Awareness subscales.

Procedure

Participants within the experimental groups completed the following phases of the study: calibration, baseline, OBL, and post-OBL. Participants in the control group did not undergo the OBL phase, instead watching a control video. After the post-OBL phase, participants from all groups completed the closing survey.

At the study's outset, participants were informed that the aim of the study was to examine physiological responses during personal pain

experiences and while observing others in pain (experimental groups), or during personal pain experiences (control group). After this, sensors measuring physiological functions were attached to the participants. Next, thermodes were placed on both forearms, and the pain calibration procedure was started.

Calibration

The temperatures of stimuli applied to each forearm were determined throughout the calibration. Half of the participants started calibration on the dominant arm, and the other half on the non-dominant arm (counterbalanced across the groups). The calibration procedure consisted of two increasing series of thermal heat stimuli applied to each arm. The first heat stimulus was 40.0 ^{o}C, and the temperature of each consecutive stimulus increased by 0.5 ^{o}C until the stimulus temperature reached 49.5 ^{o}C or the participant rated the intensity of pain elicited by the current heat stimulus as 7 or more on the NRS. The target temperatures that correspond to moderate pain (i.e., pain of intensity rated as 5 and 6 points on NRS) were calculated using an exponential trend line ($y = a^{b*x}$) and then applied in subsequent study phases.

Baseline

The 16 thermal stimuli were applied alternately to both forearms. The thermal stimuli that corresponded to pain ratings 5 and 6 on the NRS were applied in a pseudo-random order (either “5-6-6-5” (x4) or “6-5-5-6” (x4)). On which arm (dominant and non-dominant) and in what order the stimuli were applied depended on the pre-planned counterbalance.

OBL

Firstly, the ointment was applied to the participants’ forearms (the arm on which the ointment was applied was counterbalanced across groups). The ointment was applied to a rectangular area on the participant’s forearm (4x5 cm) marked with blue tape. Participants received no information on the purported effects of the ointment. If participants

asked questions, a standardized answer was provided, emphasizing that the ointment was safe and tested.

Participants from the experimental groups were notified that they would watch two videos allegedly presenting another participant performing two tasks. Participants were informed that they would subsequently be randomly assigned to undertake one of these tasks; however, in fact, participants were consistently assigned to the second task.

The videos presented a male model (34 years old) sitting in the same room as the participant. The model had the ointment applied on the same arm as the participant. The recording clearly showed half of the model's face, both arms, and a computer monitor. Participants observed the model undertaking two tasks:

As part of the accuracy manipulation, the first video presented the model with one thermode attached to the forearm without placebo ointment. The model undertook a stimuli discrimination task. He received five sequences of two stimuli and then had to choose which of the stimuli was more intense, subsequently receiving feedback on whether his judgment was correct or incorrect. When an incorrect choice was made, the model's screen displayed red text reading "incorrect," accompanied by a recorded voice stating the same. Conversely, for a correct choice the model's screen presented green text stating "correct," accompanied by a corresponding recorded voice message. Depending on which group the participants were randomized to, they viewed either the model with a high-accuracy rating or a low-accuracy rating. For the group with a model showing high accuracy, 80% of his responses were correct; for the group with a model showing low accuracy, 20% were correct. After the video, participants were asked questions regarding the model's accuracy in differentiating thermal stimuli and were asked to give predictions regarding their own accuracy (for details see "*Perceived accuracy*")

Next, participants were asked to watch the second video, which aimed to induce placebo hypoalgesia. The second video showed the same model

with two thermodes attached to his forearms. The model rated the intensity of pain elicited by thermal stimuli applied with and without the placebo; however, in fact, no painful stimuli were applied. When the model allegedly received thermal stimuli on the forearm coated with a placebo ointment, an additional icon representing ointment appeared on the computer screen. The model rated the pain elicited by stimuli applied with a placebo as low (1–3 on the NRS) and without a placebo as high (7–9 on the NRS). The model demonstrated responses to 16 stimuli applied with or without a placebo. The participants received the same sequence of stimuli during the post-OBL.

Participants in the control group watched a video showing fish in an aquarium; this video ran for the same amount of time as the videos in the OBL groups.

Post-OBL

During the post-OBL phase, the set of 16 stimuli of moderate intensity (for details see *Baseline*) was alternately applied to both forearms. When the thermal stimulus was administered to the forearm covered with a placebo ointment, an extra icon depicting ointment appeared on the computer screen.

Closing survey

To verify whether participants had been involved in observing the model, participants rated 1) how the ointment affected the model's pain and their own pain (on a scale from -5, '*decreased pain*', to 5, '*increases pain*'), and 2) how much pain the model and participants themselves experienced on average with and without the ointment (0–10 NRS). In addition, participants were asked how the model's pain rating affected their pain perception and how much they tried to adjust their pain ratings to the model's ratings. Participants were also asked to fill out a few questionnaires measuring psychological traits (for details, see *Psychological traits*).

Statistical analyses

First, descriptive statistics were calculated for age, BMI, as well as questionnaire scores; a one-way analysis of variance (ANOVAs) was conducted to detect group differences, and the chi square test was carried out to determine if the proportions of men and women as well as right-handed and left-handed participants differed between groups.

To assess if OBL can induce placebo hypoalgesia regardless of the model's accuracy (H1), a mixed-model repeated-measures ANOVA was conducted on the pain ratings with group (merged OBL groups vs. control) as a between-subjects factor, and trials (placebo, control), as well as phase (baseline, post-OBL) as within-subjects factors (2 x 2 x 2). To determine whether the ability of the model to accurately differentiate pain intensity influences the magnitude of placebo hypoalgesia (H2), a second mixed-model repeated-measures ANOVA (3 x 2 x 2) was conducted with group (high-accuracy vs. low-accuracy vs. control) as a between-subjects factor, and trials and phase as within-subjects factors. Both ANOVAs were followed by planned comparisons for differences in pain ratings: 1) between groups, in the placebo trials in the post-OBL; 2) within groups, between placebo and control trials in the post-OBL; and 3) within groups, between phases in the placebo trials. The planned comparisons were corrected for the number of comparisons (Bonferroni). Parallel analogous analyses were performed for physiological measures (HR, SCL and SCR). Square root transformation (sqr) was applied to SCR and SCL data to normalize their distribution before conducting the parametric analyses. Wherever the sphericity assumption was not met in the repeated-measures ANOVA, the Greenhouse-Geisser correction was applied. When interpreting the effect sizes, it was assumed that $\eta 2 = 0.01$ indicates a small effect, $\eta 2 = 0.06$ indicates a medium effect, and $\eta 2 = 0.14$ indicates a large effect [12]. Sensitivity analysis was conducted to evaluate the impact of participants who suspected the true aim of the study on the results.

Mediation analysis was performed to determine whether expectations mediated the effects of OBL (H3) on pain ratings in the experimental groups relative to the control group. In this analysis, group (no-observation vs high-accuracy vs low-accuracy) was treated as an independent variable; the difference in mean expectations ratings between baseline and post-OBL in placebo trials was treated as a mediator, and the difference in mean pain ratings between baseline and post-OBL in placebo trials was treated as a dependent variable. The difference in mean pain ratings between baseline and post-OBL in control trials was controlled for. Mediation analysis was preceded by a mixed-model, repeated-measures ANOVA conducted on expectations ratings to determine whether the differences in expectations mirrored those in pain ratings. Parallel analyses were conducted for physiological data. The second mediation analysis was performed to explore whether the perceived accuracy of the model mediated the effects of observational learning on pain ratings. Since the correlation between the two measures of the model's accuracy used in the study was very high ($r=0.95$, $p<0.001$), we used only one of the measures (how accurately the model estimated the difference between the stimuli) in the mediation analysis. The mediation model was constructed similarly as the first one, the difference being that in this analysis the high-accuracy group was compared to the low-accuracy group (independent variable). Bias-corrected bootstrap 95% confidence intervals (N = 5000) were used to determine if the indirect effect was significant.

The associations between interindividual traits and the magnitude of placebo hypoalgesia (operationalized as the pain decrease in placebo trials in experimental groups) induced by observing the models were explored with correlational analyses. Moderation analysis was performed to investigate whether the participants' own accuracy in pain reporting modified placebo hypoalgesia elicited by OBL. The moderator variable was operationalized as the average coefficient of variation for pain ratings during baseline. Similarly to the mediation analyses, the group

was treated as an independent variable, the difference between baseline and post-OBL pain ratings in control trials was treated as a covariate, and the change in pain ratings from baseline to post-OBL in placebo trials was treated as a dependent variable.

A one-way ANOVA followed by pairwise comparisons was conducted to analyze the answers given to the control questions (perceived accuracy and closing questions). IBM SPSS software for Windows version 28 (IBM Corp., Armonk, NY, USA) was used for all of the analyses. Mediation and moderation analyses were performed using the Process Macro version 4.1 [22], applying models number 4 (mediation) and number 1 (moderation). The alpha level was set at 0.05 (two-sided tests) for rejection of the null hypothesis in all the statistical analyses.

Results

Preliminary analyses

Data was collected from 130 participants. 34 participants were excluded from the study due to a modification in the calibration method that was implemented after they were tested. The final sample consisted of 96 participants (60 women, 36 men). There were no significant differences between groups in age, BMI, LOT-R, EQ-Short, IRI, FPQ, SPSRQ and BAIAS (all $p \geq 0.10$). Moreover, the proportions of men/women and right-handed/left-handed participants did not differ across groups (all $p \geq 0.67$). Details are presented in Table 1.

Table 1. Descriptive statistics and differences between groups

Variable	Mean (SD)				Test statistic	p
	Total sample (N=96)	No observation (N=32)	High accuracy (N=32)	Low accuracy (N=32)		
Age	22.75(4.89)	23.31(6.65)	22.81(3.69)	22.13 (4.04)	F(2,88)=0.43	0.66
BMI	22.10(3.5	22.19(3.1	22.48(4.33	21.62(3.00	F(2,88)=0.	0.64

	5)	6)	)	)	46	
LOT-R	14.74(4.88)	15.50(5.29)	14.53(4.77)	14.19(4.64)	F(2,93)=0.62	0.54
EQ-Short	23.62(6.77)	23.41(6.30)	23.34(7.28)	24.09(6.87)	F(2.93)=0.12	0.89
IRI Empathic Concern	30.23(7.18)	31.25(6.18)	29.31(8.74)	30.13(6.46)	F(2,93)=0.58	0.56
IRI Personal Distress	17.04(5.40)	16.34(6.24)	16.31(5.40)	18.47(4.27)	F(2,93)=1.70	0.19
IRI Perspective Taking	25.54(5.16)	25.22(5.14)	24.94(5.45)	26.47(4.90)	F(2,93)=0.80	0.45
FPQ Total	73.81(15.47)	76.84(14.82)	71.86(15.07)	72.57(16.51)	F(2,87)=0.92	0.40
FPQ Severe Pain	31.98(8.19)	33.78(7.70)	32.52(8.24)	29.69(8.33)	F(2,90)=2.14	0.12
FPQ Minor Pain	18.37(5.29)	18.84(5.33)	18.28(5.79)	17.97(4.83)	F(2,92)=0.22	0.80
FPQ Medical Pain	23.96(6.51)	24.19(6.81)	22.72(5.60)	25.00(7.07)	F(2,91)=1.00	0.37
SPSRQ Sensitivity to Rewards	4.98(2.10)	4.38(2.18)	5.50(1.88)	5.06(2.14)	F(2,93)=2.40	0.10
SPSRQ Sensitivity to Punishments	6.18(3.43)	5.50(3.66)	6.38(3.17)	6.66(3.43)	F(2,93)=0.99	0.38
BAIAS Body Ignorance	1.93(0.59)	1.82(0.57)	2.04(0.63)	1.92(0.57)	F(2,93)=1.10	0.34
BAIAS Body Awareness	3.26(0.57)	3.34(0.54)	3.37(0.54)	3.09(0.62)	F(2,93)=2.38	0.10
BAIAS Body Attention	2.62(0.66)	2.58(0.60)	2.75(0.74)	2.52(0.63)	F(2,93)=1.05	0.35
			Number of participants			

Sex (Women/Men)	54/43	21/11	21/11	18/14	$X^2(2)=0.80$	0.67
Dominant hand (Right-handed/ Left-handed)	88/8	29/3	29/3	30/2	$X^2(2)=0.27$	0.87

Note: LOT-R – Life Orientation Test-Revised [44], EQ-Short – Empathy Quotient [56], IRI – Interpersonal Reactivity Index [20], FPQ – Fear of Pain Questionnaire [34], SPSRQ – Sensitivity to Punishment and Reward Questionnaire [16], BAIAS – Body Attention, Ignorance and Awareness Scale [7].

Primary analyses

Effects of OBL on pain ratings

The three-way group (merged experimental groups vs no-observation group) x trial x phase interaction effect was statistically significant [F (1, 94) = 4.24, p=0.042, ηp2 = 0.04], indicating that, as hypothesized (H1), observational learning (OBL) occurred. The planned comparisons showed that the OBL groups and the no-observation group did not differ significantly in terms of the mean pain experienced in the placebo trials in post-OBL (p=0.85). Interestingly, the participants from the OBL groups reported a borderline significant increase in baseline to post-OBL pain ratings for placebo trials (p=0.057). Additionally, further exploratory analyses (corrected for all pairwise comparisons) revealed that pain also increased significantly from baseline to post-OBL for control trials in the OBL groups (p<0.001). The difference between pain ratings for placebo and control trials (measured post-OBL) turned out to be significant only within the OBL groups: after observing the model, participants from these groups experienced less pain in the placebo trials relative to the control trials (p=0.012). Estimated marginal means for pain ratings are presented in Figure 2.

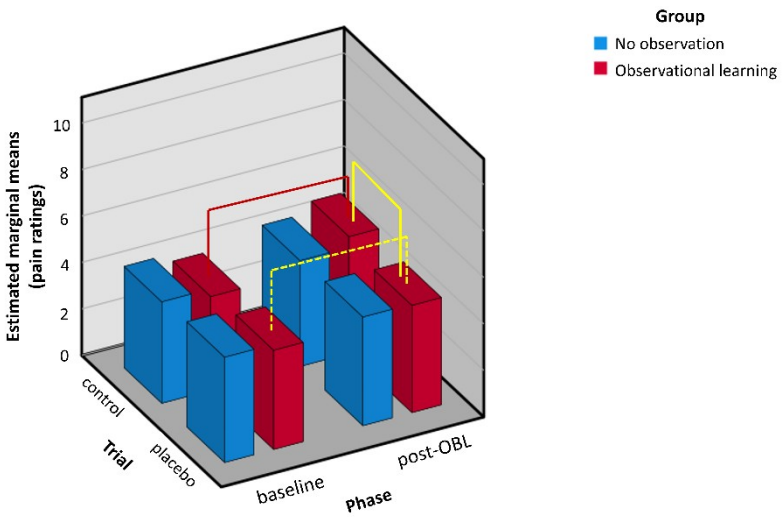

Figure 2. Estimated marginal means for pain ratings during application of the thermal stimuli in the different groups (no-observation, observational learning), phases (baseline, post-OBL) and trials (placebo, control) of the experiment. *Note: significant differences between means are represented by colored lines (Bonferroni-adjusted significance tests for pairwise comparisons): yellow line p<0.05, red line p<0.001, yellow dotted line p=0.06*

Effects of model accuracy on OBL-induced effects for pain ratings

When the two OBL groups were separated in order to assess the impact of model accuracy on the magnitude of OBL-induced effects in pain (H2), no three-way interaction effect was found [F (2, 93) = 2.12, p=0.126, ηp2 = 0.04]. Planned comparisons revealed no difference between the low- and high-accuracy groups (p=0.17). However, they showed that the difference between pain ratings in placebo and control trials after watching the model was significant only in the high-accuracy group (p=0.008), indicating that only these participants experienced lower pain in the placebo trials than in the control trials. See Figure 3 for details.

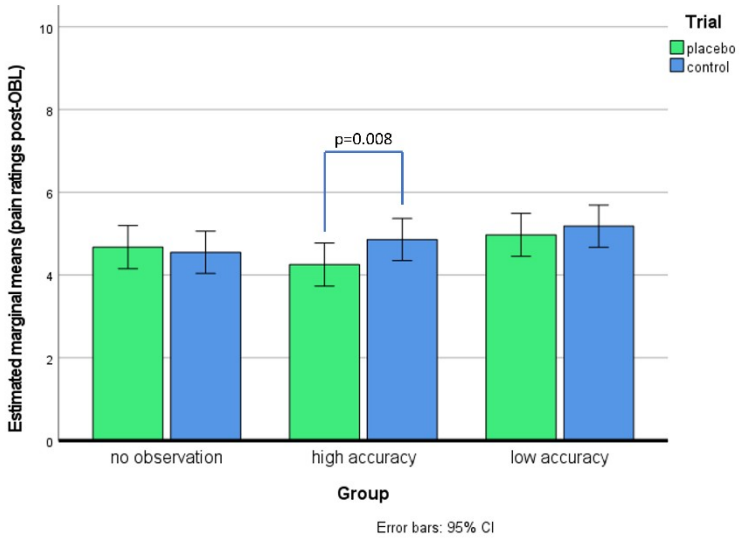

Figure 3. Estimated marginal means for pain ratings during application of the thermal stimuli in the different groups (no-observation, observational learning), and trials (placebo, control) in the post-OBL. *Note: significant differences between means are represented by colored lines (Bonferroni-adjusted significance tests for planned comparisons).*

Sensitivity analyses

The estimates did not change substantially after excluding the participants who had suspicions about the aim of the study (N=3 from the high-accuracy group, and N=4 from the low-accuracy group). The three-way interaction effect of group (OBL, no-observation) x trial (placebo, control) x phase (baseline, post-OBL) remained significant [F (1, 87) = 4.44, p=0.038, ηp2 = 0.05] and became nonsignificant when the OBL groups were further separated to evaluate the effects of model accuracy: F (2, 86) = 2.35, p=0.101, ηp2 = 0.05. Planned comparisons also revealed analogous results.

Secondary analyses

Effects of OBL on physiological outcomes

A mixed-model ANOVA performed on physiological outcomes (HR, $SCR_{sqr,}$ SCL_{sqr}) showed that the hypothesized group (OBL vs no-observation) x trial x phase interaction effects were not significant (all p> .46). Planned

comparisons revealed no significant differences between groups in the placebo trials in post-OBL for all physiological measures (all p>.09). Differences in SCR_{sqr} between placebo and control trials (post-OBL) were significant in the OBL group (M_{diff}= -.07, p<.001) but not in the control group (M_{diff}= -.02, p=.439). On the other hand, between-trial differences in HR and SCL_{sqr} were non-significant in both groups (all p > .06). SCR_{sqr} (M_{diff}=.09, p<.001) and SCL_{sqr} (M_{diff}=.12, p<.001) significantly decreased from baseline to post-OBL (placebo trials) in the OBL group but not in the no-observation group (all p>.34), and HR significantly decreased in both the OBL (M_{diff}=2.23, p<.001) and the no-observation groups (M_{diff}=3.30, p<.001).

When the OBL groups were divided according to the accuracy of the model, the three-way group (high-accuracy, low-accuracy, no-observation) x trial x phase interaction effects were not significant for any physiological outcome (all p > .56). The results of the planned comparison mostly aligned with those from the merged OBL groups analyses. SCR_{sqr} differences between control trials and placebo trials (post-OBL) were significant only in the high-accuracy group (M_{diff}= -.09, p=.002). The SCR_{sqr} decline from baseline to post-OBL in placebo trials was significant only in the low-accuracy group (M_{diff} = .12, p=.001). Both the high- (M_{diff}=.12, p=.005) and low-accuracy groups (M_{diff}=.12, p=.004) showed SCL_{sqr} decreases from baseline to post-OBL. HR decreases were observed across all groups: high-accuracy group (M_{diff}=2.12, p<.001), low-accuracy group (M_{diff}=2.34, p<.001), and no-observation group (M_{diff}=3.30, p<.001).

Expectations as a mediator of the response to a placebo

A mixed-model repeated-measures ANOVA with group (high- vs low-accuracy vs no-observation) as a between-subjects factor and trial and phase (baseline, post-observation 1/ post-OBL1, post-observation 2/ post-OBL 2) as within-subject factors revealed a significant three-way interaction [F(3.37,156.55) = 9.40, p < .001, ηp2 = .17], showing that

model observation influenced pain expectations. After observing the model, the high-accuracy group anticipated less pain in placebo trials than in control trials. Notably, some of the expectation patterns differed from the pain ratings. Specifically, the no-observation group and the low-accuracy group differed significantly in expectation ratings during control trials in the post-OBL1 phase (p=0.018), with the low-accuracy group expecting more pain. Within the low-accuracy group, a significant reduction in expected pain from baseline to post-OBL1 in placebo trials was noted (p<0.001). Also there was a difference between placebo and control trials in post-OBL1 for both high and low-accuracy groups (p<0.001). For the full results, see Supplementary materials (Figure S1).

The indirect effect on pain ratings via expectations. The mediation analysis indicated that total effects of group on change in pain ratings from baseline to the post-OBL in placebo trials was nonsignificant for both the high-accuracy vs. no-observation group (b=-0.27, SE=0.33, LCI95%=-0.2, UCI95%=0.39; p=0.42), and for the low-accuracy vs. no-observation group (b=-0.28, SE=0.33, LCI95%=-0.94, UCI95%=0.39; p=0.41). Nevertheless, the indirect effect of the low-accuracy vs. no-observation group on the dependent variable was statistically significant (b=-0.25, SE=0.15, LCI95%=-0.59, UCI95%=-0.02). (H3) (Figure 4).

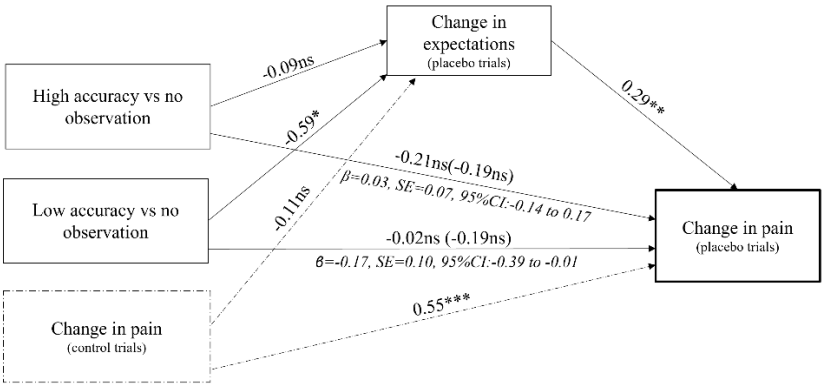

Figure 4. Results of the analysis testing the mediating role of expectations on the change in pain ratings. *Note: the change in expectations was operationalized as the difference between post-OBL and*

*baseline expectations ratings in placebo trials; the change in pain was operationalized as the difference between post-OBL and baseline pain ratings in placebo trials; total and direct effects are insignificant; the controlled variable is represented by the dotted lines; total effect is presented in brackets – R^2 for the model =0.32, F(4,91)=10.72, p<0.001; *p<0.05, **p<0.01, ***p<0.001; ns – not significant.*

The indirect effects via expectations on physiological measures. The indirect effects of group (high-accuracy vs no-observation, and low-accuracy vs no-observation) on change in physiological outcomes (HR, SCR_{sqr}, SCL_{sqr}) via change in expectations were not significant. For details, see Supplementary materials (Figures S2–S4).

Perceived accuracy of the model as a mediator of the response to a placebo

Total effect of low accuracy vs high accuracy on change in pain ratings was close to zero and not significant (b=0.003, SE=0.35, LCI95%=-0.70, UCI95%=0.70; p=0.99). Although the groups differed significantly in their estimations of model accuracy (b=-50.33, SE=2.82, LCI95%=-55.96, UCI95%=-44.69; p<0.001), and the difference was in the assumed direction, there was no association between the model's accuracy and the change in pain ratings in placebo trials (b=0.02, SE=0.02, LCI95%=-0.02, UCI95%=0.05; p=0.31). Not surprisingly, the indirect effect of group via perception of accuracy on the dependent variable was also not significant (b=-0.82, SE=0.62, LCI95%=-2.15, UCI95%=0.30). See Fig. 5 for details.

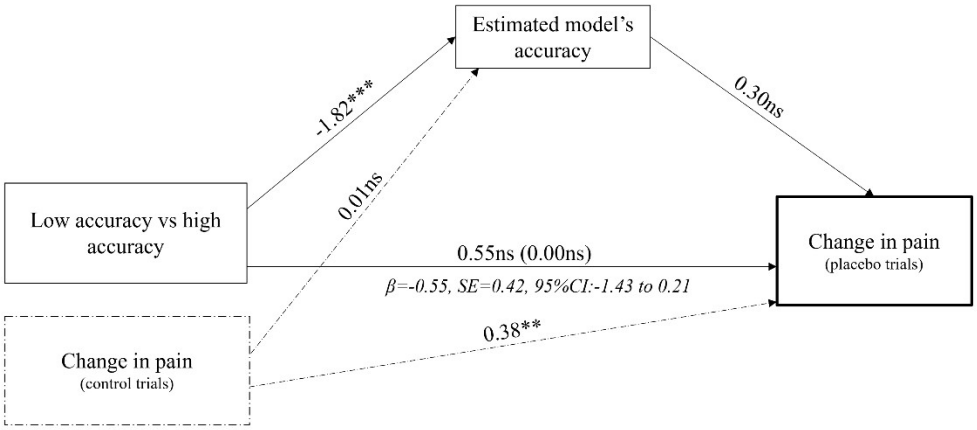

Figure 5. Results of the analysis testing the mediating role of the model's estimated accuracy in response to a placebo. *Note: standardized estimates are presented; total effect, direct effect and indirect effect are not significant; the controlled variable is represented by dotted lines; total effects are presented in brackets; R^2 for the model =0.16, F(3,60)=3.81, p=0.014; **p<0.01, ***p<0.001.*

Exploratory analyses

Associations between individual characteristics and changes in pain ratings

The correlation analyses did not yield any significant results. Response to a placebo was not associated with any of the participants' individual characteristics (all p>0.16). For details, see Supplementary materials (Table S1).

Variance in pain reporting as a moderator of the response to a placebo

The average CoV was calculated as the mean of four CoVs: one calculated for each arm at each stimulus level at baseline [52]. The interaction effect of group and average CoV on change in pain ratings from baseline to post-OBL in placebo trials was not significant, b=-0.22, SE=1.10, LCI95%=-2.41, UCI95%=1.97, p=0.84, indicating that variance

in pain ratings does not moderate response to a placebo. Estimates of the full model can be found in Supplementary materials (Table S2).

Control questions: perceived accuracy and closing survey

The high-accuracy group rated the model's accuracy as significantly higher (EMM=72.18, SE=1.97) than the low-accuracy group (EMM=21.91, SE=1.97), $F(1,62)=324.64$, $p<0.001$, $\eta p2=0.84$. Similarly, they rated their own accuracy as higher (EMM=66.77, SE=2.50) compared to the low-accuracy group (EMM=42.60, SE=2.50), $F(1,62)=46.80$, $p<0.001$, $\eta p2=0.43$.

When asked about the intensity of the pain that was experienced during heat stimuli on the arm without ointment in the closing survey, participants from both OBL groups declared that they felt significantly higher pain (EMM=5.55, SE=0.17) than participants from the no-observation group (EMM=4.84, SE=0.24): $F(1,94)=5.57$, $p=0.020$, $\eta p2=0.06$. The high- and low-accuracy OBL groups did not differ on any of the other items of the closing survey (all $p > 10$). Details can be found in Supplementary materials (Table S3).

Discussion

One of the study aims was to induce placebo hypoalgesia through OBL. However, participants' responses to stimuli applied with placebo did not decrease due to OBL, and even a slight non-significant increase in pain was observed from baseline to post-OBL. Instead, participants learned to respond with hyperalgesia to stimuli applied without placebo. These data suggest that observers' pain responses were shaped mainly by the hyperalgesic reactions of the observed model that were demonstrated in the non-placebo condition. Similar responses to placebo have also been observed in previous studies [35,50]. In these studies, participants reacted with higher pain to the cues associated with the model's more painful responses, but their reactions to pain stimuli preceded by cues associated with the model's hypoalgesic responses did not change from

the baseline to post-OBL [35] or did not differ from the reactions of participants who were not subjected to OBL [50].

Observer's attention is a critical process of OBL [6], and could explain these results. Attention can be biased toward predictors of aversive outcomes [19,24,46]. In the current experiment, participants observed the model's reactions in two conditions that differed in the level of aversiveness: in the non-placebo condition the model experienced more intense pain than in the placebo condition. It seems that observers' attention was attracted mainly by the contextual cues signaling the absence of placebo (for example, a thermometer icon displayed alone), namely the cues associated with the more aversive condition. As a result, participants learned to respond with greater pain in the absence of the placebo, but they did not learn to respond to the placebo. As the participants felt more pain in a context that had previously accompanied the model's greater pain, it can be assumed that a kind of nocebo hyperalgesia rather than placebo hypoalgesia was elicited in the observers. Since participants in studies on observationally induced placebo effects are usually exposed simultaneously to the model's analgesic and hyperalgesic responses (and thus the cues associated with hyperalgesic responses) [5,8,9,14,23,41,43,45,50,51,53–55], establishing the involvement of OBL in shaping placebo hypoalgesia appears difficult [37].

The psychophysiological responses measured during the experiment decreased over time, which seems to result from acclimatization to the experimental environment rather than from the experimental manipulation. As no clear physiological markers of pain have been identified to date [18], it is possible that the physiological changes observed in participants reflected changes in not so much their pain experiences but in their stress levels, or possibly changes in physical activity after traveling to the lab.

The existing studies suggest that the individual characteristics of the observed person contribute to placebo effects in pain. Participants' perception of the model's social status [8] and self-confidence [9] have been found to be predictors of placebo hypoalgesia. However, observing a male model demonstrating pain significantly increased the magnitude of nocebo hyperalgesia compared to observing a female model [50]. This indicates that certain model's characteristics, such as sex, have the potential to substantially alter placebo effects. In the current study, only the model perceived as having good pain assessment skills was effective in shaping nocebo hyperalgesia. Thus, it seems that social learning of placebo effects in pain is governed mainly by the characteristics of the model that allow pain predictions. It has been shown that gender-related stereotypes can significantly modulate pain perception [39,48,58]. According to these stereotypes, males are perceived as less sensitive to pain and less willing to demonstrate pain compared to females. Such a bias might cause a male model's high pain ratings to be perceived as more threatening to the participants themselves. Thus, the sex of the model experiencing pain may be an easily accessible cue that facilitates inferences about the intensity of upcoming pain. Similarly, confidence in the demonstrator's ability to assess pain accurately may increase the reliability of their pain judgments, thus making them good predictors of an aversive event. On the other hand, knowledge about the model's self-confidence or social status, although related to social learning [10], may be less helpful in interpreting observed pain and predicting upcoming pain. Future research could focus more on the characteristics of models that could influence the communication of painful states (e.g., fear of pain or pain tolerance).

In the current study, OBL influenced participants' expectations, which is in line with the theoretical model of placebo effects in pain [2,3] and the results of a meta-analysis [37]. Interestingly, expectations in participants who observed high- and low-accuracy models developed differently. Participants observing a highly accurate model expected more pain in the

post-OBL phase compared to the baseline phase when thermal stimuli were applied without a placebo. OBL did not influence participants' expectations related to stimuli applied with a placebo; thus, their expectations aligned with their pain responses to stimuli administered with and without a placebo. Participants observing the low-accuracy model also expected more pain in the post-OBL phase when thermal stimuli were applied without a placebo. However, their expectations related to stimuli applied with a placebo were lower than their baseline expectations. Since the low-accuracy model was not perceived as a reliable source of information regarding impending pain, observers might have attended not so much to the model's hyperalgesic responses but to the whole context accompanying all the observed pain-related behaviors, thus developing pain-related expectations on this basis. The pattern of expectations observed in this group could potentially lead to placebo hypoalgesia. Based on this result, we speculate that the features of the model may not only influence the magnitude of placebo effects induced by OBL [50], but could also be considered a factor that determines the type of effect elicited through OBL.

Pain-related expectations were not involved in nocebo hyperalgesia induced in participants observing a highly accurate model, which is in line with data showing that cognitive processes may not necessarily mediate nocebo [1] or placebo effects [13]. Interestingly, expectations mediated pain responses in participants observing a low-accuracy model. This result aligns with data showing that the mediating role of expectations may be particularly evident when available information concerning pain is incongruent [4]. It seems that information on pain received from a source perceived as low accuracy could work similarly to incongruent information on pain.

The current study also investigated the associations between observers' characteristics and placebo effects. Empathy, although linked to social learning of placebo effects [2], was not associated with participants' ability to internalize the experiences of the observed person. This result

aligns with a meta-analysis showing that empathy is relatively weakly related to placebo effects induced by OBL [37]. Prior research identified optimism [28,38], fear of pain [21], sensitivity to bodily signals [36], and reward responsiveness [11,36] as predictors of placebo effects. The current study did not confirm these associations, which aligns with the results of recently published meta-analysis [27]. However, it should be noted that the sample sizes of the studies cited above and those included in the meta-analysis could be insufficient to detect associations between observers' individual differences and their responses to a placebo.

One limitation of this study is the healthy, young sample; therefore, these findings should be generalized to the clinical population with caution. In this study, as in previous ones, the observer was simultaneously exposed to the hyperalgesic and hypoalgesic responses of the model. This procedure makes it difficult to determine the role of OBL in developing hypoalgesic reactions to placebo in an observer. To eliminate this issue, future studies may consider studying OBL with adapted experimental paradigms. For instance, it would be advisable to administer pain stimuli to the participant, then present a placebo-treated model who responds to these pain stimuli with lesser pain than the participant, after which the participant's responses to the same pain stimuli applied with and without a placebo would be checked. A limitation of this study could also be the design of the control group. Although participants in this group did not observe a model experiencing pain, they watched a video of swimming fish, which may have relaxed them or distracted them from the pain, thus inadvertently affecting their pain responses.

Some methodological advantages should also be noticed. The efficacy of the manipulation of model's accuracy was controlled for. Data showed that participants from the experimental groups differed in their estimation of the observed model's accuracy, and these differences were in line with the assumed directions. However, the estimated accuracy of the model did not mediate participants' responses to pain. Observing high- and low-accuracy models also influenced participants' predictions

regarding their accuracy in differentiating stimuli of different intensities. However, participants' actual accuracy in pain reporting during the baseline did not moderate the effects elicited by OBL. To control for the participants' level of involvement in the experiment, a series of questions were asked at the end of the experiment.

The findings from this study might have practical implication. Considering the efficacy of highly accurate models in shaping responses to placebo, it would be advisable to involve well-adjusted patients as models to demonstrate the potentially positive effects of treatment.

Acknowledgments

The work described in the manuscript was funded by the National Science Centre in Poland (grant no. 2019/35/B/HS6/04320, granted to Przemysław Bąbel). The authors have no conflict of interest to declare.

References

[1] Bąbel P, Adamczyk W, Świder K, Bajcar EA, Kicman P, Lisińska N. How Classical Conditioning Shapes Placebo Analgesia: Hidden versus Open Conditioning. Pain Med 2018;19:1156–1169.

[2] Bajcar EA, Bąbel P. How does observational learning produce placebo effects? A model integrating research findings. Front Psychol 2018;9:2041.

[3] Bajcar EA, Bąbel P. Social Learning of Placebo Effects in Pain: A Critical Review of the Literature and a Proposed Revised Model. J Pain 2024:104585.

[4] Bajcar EA, Wiercioch-Kuzianik K, Adamczyk WM, Bąbel P. To Experience or to Be Informed? Classical Conditioning Induces Nocebo Hyperalgesia even when Placebo Analgesia Is Verbally Suggested-Results of a Preliminary Study. Pain Med 2020;21:548–560.

[5] Bajcar EA, Wiercioch-Kuzianik K, Farley D, Adamczyk WM, Buglewicz E, Bąbel P. One of us or one of them? The effects of the model's and observer's characteristics on placebo analgesia induced by observational learning. PLOS ONE 2020;15:e0243996.

[6] Bandura A. Social Learning Theory. 1st edition. Upper Saddle River: Prentice-Hall, 1976.

[7] van Beugen S, Ograczyk A, Ferwerda M, Smit JV, Zeeuwen-Franssen MEJ, Kroft EBM, de Jong EMGJ, Zalewska-Janowska A, Donders AR, van de Kerkhof PCM, van Middendorp H, Evers AWM. Body attention, ignorance and awareness scale: assessing relevant concepts for physical and psychological functioning in psoriasis. Acta Derm Venereol 2015;95:444–450.

[8] Bieniek H, Bąbel P. The Effect of the Model's Social Status on Placebo Analgesia Induced by Social Observational Learning. Pain Med 2022;23:81–88.

[9] Brączyk J, Bąbel P. The Role of the Observers' Perception of a Model's Self-Confidence in Observationally Induced Placebo Analgesia. J Pain 2021;22:1672–1680.

[10] Brewer KR, Wann DL. Observational learning effectiveness as a function of model characteristics: investigating the importance of social power. 1998;26(1):1–10.

[11] Broelz EK, Enck P, Niess AM, Schneeweiss P, Wolf S, Weimer K. The neurobiology of placebo effects in sports: EEG frontal alpha asymmetry increases in response to a placebo ergogenic aid. Sci Rep 2019;9:2381.

[12] Cohen. Applied multiple regression/correlation analysis for the behavioral sciences. 2nd ed. Hillsdale, NY: Lawrence Erlbaum Associates, 1983.

[13] Colloca L, Akintola T, Haycock NR, Blasini M, Thomas S, Phillips J, Corsi N, Schenk LA, Wang Y. Prior Therapeutic Experiences, Not Expectation

Ratings, Predict Placebo Effects: An Experimental Study in Chronic Pain and Healthy Participants. Psychother Psychosom 2020;89:371–378.

[14] Colloca L, Benedetti F. Placebo analgesia induced by social observational learning. Pain 2009;144:28–34.

[15] Colloca L, Miller FG. How placebo responses are formed: a learning perspective. Philos Trans R Soc Lond, B, Biol Sci 2011;366:1859–1869.

[16] Cooper A, Gomez R. The development of a short form of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire. Journal of Individual Differences 2008;29:90–104.

[17] Corsi N, Colloca L. Placebo and Nocebo Effects: The Advantage of Measuring Expectations and Psychological Factors. Front Psychol 2017;8:308.

[18] Cowen R, Stasiowska MK, Laycock H, Bantel C. Assessing pain objectively: the use of physiological markers. Anaesthesia 2015;70:828–847.

[19] Crombez G, Van Ryckeghem DML, Eccleston C, Van Damme S. Attentional bias to pain-related information: A meta-analysis. PAIN 2013;154:497–510.

[20] Davis, M. H. A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology 1980;10:85.

[21] Geers AL, Faasse K, Guevarra DA, Clemens KS, Helfer SG, Colagiuri B. Affect and emotions in placebo and nocebo effects: What do we know so far? Social and Personality Psychology Compass 2021;15.

[22] Hayes AF. Mediation, moderation, and conditional process analysis. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York: Guilford Publications, 2017.

[23] Hunter T, Siess F, Colloca L. Socially induced placebo analgesia: a comparison of a pre-recorded versus live face-to-face observation. European Journal of Pain 2014;18:914–922.

[24] Jackson T, Su L, Wang Y. Effects of Higher Versus Lower Threat Contexts on Pain-Related Visual Attention Biases: An Eye-Tracking Study of Chronic Pain. J Pain 2018;19:649–659.

[25] Jankowiak-Siuda K, Kantor-Martynuska J, Siwy-Hudowska A, Śmieja M, Dobrołowicz-Konkol M, Zaraś-Wieczorek I, Siedler A. Analiza właściwości psychometrycznych polskiej wersji językowej Skróconej Skali Ilorazu Empatii (SSIE) – The Empathy Quotient (EQ-Short). Psychiatr Pol 2017;51:719–734.

[26] Juczyński, Poprawa. The Life Orientation Test – Revised (LOT-R). Polish adaptation. In: Zygfryd Juczyński, editor. Narzędzia pomiaru w promocji i psychologii zdrowia. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego. Warszawa: Pracownia Testów Psychologicznych PTP, 2001. pp. 60–70.

[27] Kang H, Miksche MS, Ellingsen D-M. Association between personality traits and placebo effects: a preregistered systematic review and meta-analysis. Pain 2023;164:494–508.

[28] Karacaoglu M, Meijer S, Peerdeman KJ, Dusseldorp E, Jensen KB, Veldhuijzen DS, van Middendorp H, Evers AWM. Susceptibility to Nocebo Hyperalgesia, Dispositional Optimism, and Trait Anxiety as Predictors of Nocebo Hyperalgesia Reduction. Clin J Pain 2023;39:259–269.

[29] Kaźmierczak M, Plopa M, Retowski S. Skala Wrażliwości Empatycznej. Przegląd Psychologiczny 2007;50(1):9–24.

[30] Kern A, Kramm C, Witt CM, Barth J. The influence of personality traits on the placebo/nocebo response: A systematic review. Journal of Psychosomatic Research 2020;128:109866.

[31] Lakens D, Caldwell AR. Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. Advances in Methods and Practices in Psychological Science 2021;4:2515245920951503.

[32] Lee V, Rutherford MD. Sixteen-month-old infants are sensitive to competence in third-party observational learning. Infant Behav Dev 2018;52:114–120.

[33] Łuszczyńska A, Cieślak R. Lęk przed bólem. Adaptacja kwestionariusza FPQ-III. Wyniki wstępnych badań nad rzetelnością i trafnością. [Fear of pain: Polish version of FPQ-III questionnaire and preliminary results on its psychometric properties.]. Studia Psychologiczne 2005;43:75–82.

[34] McNeil DW, Rainwater AJ. Development of the Fear of Pain Questionnaire-III. J Behav Med 1998;21:389–410.

[35] Meeuwis SH, Kłosowska J, Bajcar EA, Wasylewski MT, Badzińska J, Rubanets D, Di Nardo M, Mazzoni G, Bąbel P. Placebo Hypoalgesia and Nocebo Hyperalgesia Induced by Observational Learning May Be Difficult to Disentangle in a Laboratory Setting. The Journal of Pain 2024;25:805–818.

[36] Meeuwis SH, van Middendorp H, Veldhuijzen DS, Evers AWM. Associations Between Interindividual Differences, Expectations and Placebo and Nocebo Effects in Itch. Front Psychol 2021;12:781521.

[37] Meeuwis SH, Wasylewski MT, Bajcar EA, Bieniek H, Adamczyk WM, Honcharova S, Di Nardo M, Mazzoni G, Bąbel P. Learning pain from others: a systematic review and meta-analysis of studies on placebo hypoalgesia and nocebo hyperalgesia induced by observational learning. Pain 2023;164:2383–2396.

[38] Morton DL, Watson A, El-Deredy W, Jones AKP. Reproducibility of placebo analgesia: Effect of dispositional optimism. Pain 2009;146:194–198.

[39] Paganini GA, Summers KM, Ten Brinke L, Lloyd EP. Women exaggerate, men downplay: Gendered endorsement of emotional dramatization stereotypes contributes to gender bias in pain expectations. J Exp Soc Psychol 2023;109:104520.

[40] Raedts,Mariet, Rijlaarsdam,Gert, van Waes,Luuk, Daems,Frans. Observational learning through video-based models: Impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performances. In: Suzanne Hidi, Pietro Boscolo, editors. Writing and Motivation. Studies in Writing. Brill, 2006. pp. 203–217.

[41] Raghuraman N, Wang Y, Schenk LA, Furman AJ, Tricou C, Seminowicz DA, Colloca L. Neural and behavioral changes driven by observationally-induced hypoalgesia. Scientific Reports 2019;9:19760.

[42] Rooney T, Sharpe L, Todd J, Tang B, Colagiuri B. The nocebo effect across health outcomes: A systematic review and meta-analysis. Health Psychol 2024;43:41–57.

[43] Rubanets D, Badzińska J, Kłosowska J, Bąbel P, Bajcar EA. Pain Rating is Worth a Thousand Words: Nocebo Hyperalgesia Induced by Verbal Modeling Prevails Over the Effects of Symbolic Modeling and Verbal Suggestion. The Journal of Pain 2024;25:104442.

[44] Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. Journal of Personality and Social Psychology 1994;67:1063–1078.

[45] Schenk LA, Colloca L. The neural processes of acquiring placebo effects through observation. NeuroImage 2020;209:116510.

[46] Schmidt LJ, Belopolsky AV, Theeuwes J. Attentional capture by signals of threat. Cognition and Emotion 2015;29:687–694.

[47] Schunk DH, Zimmerman BJ. Social origins of self-regulatory competence. Educational Psychologist 1997;32:195–208.

[48] Schwarz KA, Sprenger C, Hidalgo P, Pfister R, Diekhof EK, Büchel C. How Stereotypes Affect Pain. Sci Rep 2019;9:8626.

[49] Selbing I, Olsson A. Beliefs about Others' Abilities Alter Learning from Observation. Sci Rep 2017;7:16173.

[50] Świder K, Bąbel P. The effect of the sex of a model on nocebo hyperalgesia induced by social observational learning. PAIN® 2013;154:1312–1317.

[51] Świder K, Bąbel P. The effect of the type and colour of placebo stimuli on placebo effects induced by observational learning. PLOS ONE 2016;11:e0158363.

[52] Treister R, Eaton TA, Trudeau JJ, Elder H, Katz NP. Development and preliminary validation of the focused analgesia selection test to identify accurate pain reporters. J Pain Res 2017;10:319–326.

[53] Tu Y, Park J, Ahlfors SP, Khan S, Egorova N, Lang C, Cao J, Kong J. A neural mechanism of direct and observational conditioning for placebo and nocebo responses. Neuroimage 2019;184:954–963.

[54] Vögtle E, Barke A, Kröner-Herwig B. Nocebo hyperalgesia induced by social observational learning. Pain 2013;154:1427–1433.

[55] Vögtle E, Kröner-Herwig B, Barke A. Nocebo hyperalgesia: contributions of social observation and body-related cognitive styles. J Pain Res 2016;9:241–249.

[56] Wakabayashi A, Baron-Cohen S, Wheelwright S, Goldenfeld N, Delaney J, Fine D, Smith R, Weil L. Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). Personality and Individual Differences 2006;41:929–940.

[57] Wytykowska A, Białaszek W, Ostaszewski P. Psychometryczne właściwości polskiej wersji krótkiej skali wrażliwości na kary i nagrody (SPSRQ-SF Cooper I Gomez, 2008). [Psychometric parameters of the Polish Short Version of Sensitivity to Punishment and Reward Scale (SPSRQ-SF Cooper I Gomez, 2008).]. Studia Psychologiczne 2014;52:28–39.

[58] Zhang L, Losin EAR, Ashar YK, Koban L, Wager TD. Gender Biases in Estimation of Others' Pain. J Pain 2021;22:1048–1059.

Supplementary material

Supplementary Figures

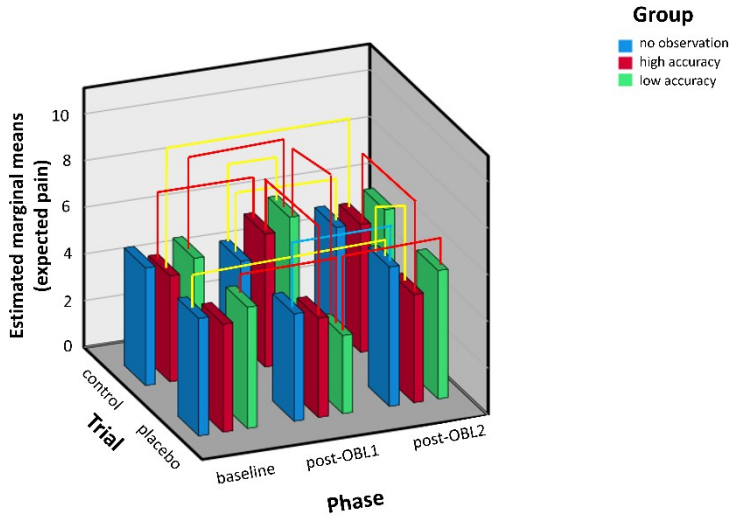

Figure S1. Estimated marginal means for expectation ratings during application of the thermal stimuli in the different groups (high-accuracy, low-accuracy, no-observation), phases (baseline, post-observation: post-OBL1, post-observation, post-OBL2) and trials (placebo, control) of the experiment. *Note: significant differences between means are represented by the coloured lines (Bonferroni-adjusted significance tests for pairwise comparisons): yellow line p<0.05, blue line p<0.01, red line p<0.001.*

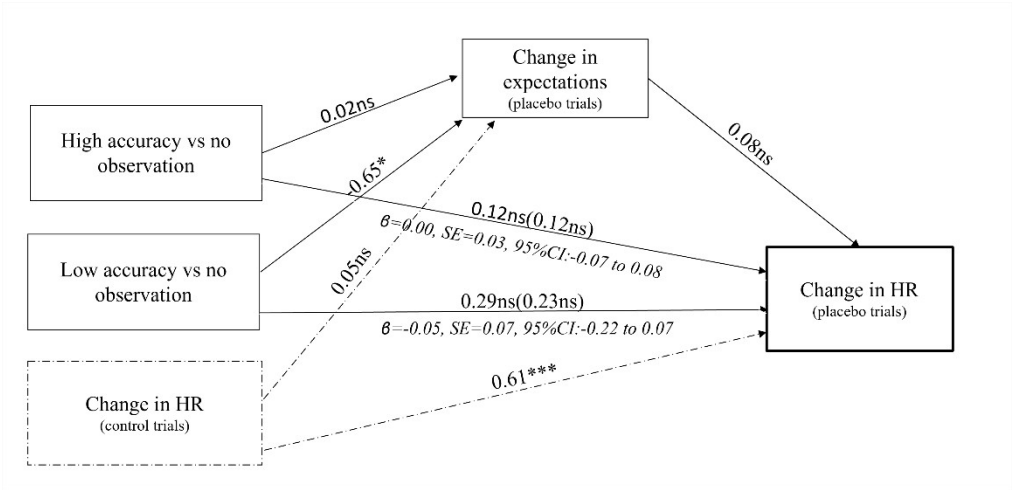

Figure S2. **Results of analysis testing the mediating role of change in expectations in between-group differences in HR change.** *Note: change in expectations was operationalized as the difference between post-OBL and baseline expectation ratings in placebo trials; change in HR was operationalized as the difference between post-OBL and baseline HR in placebo trials; total, direct and indirect effects are not significant; the controlled variable is represented by dotted lines; total effect is presented in brackets;* R^2 *for the model =0.40, F(4,81)=13.41, p<0.001; *p<0.05, ***p<0.001; ns – not significant.*

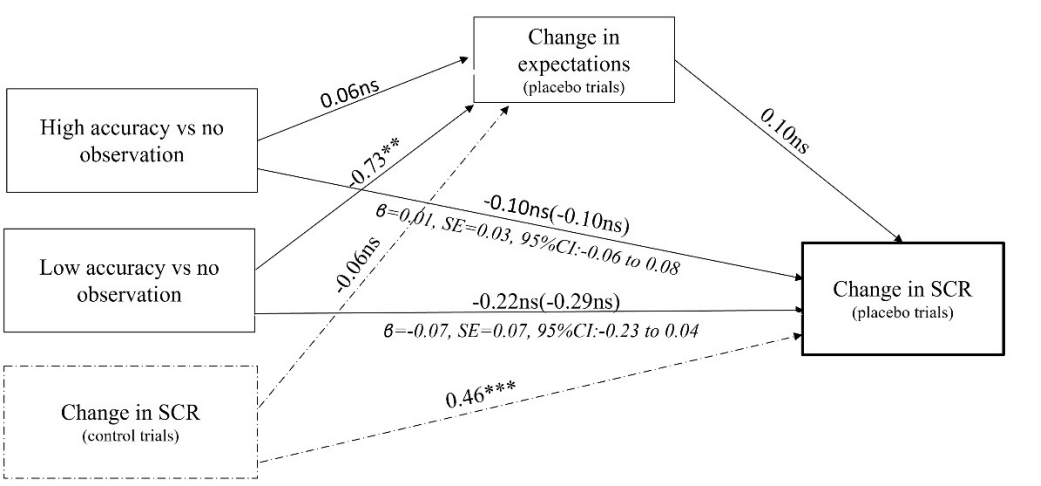

Figure S3. Results of analysis testing the mediating role of change in expectations in between-group differences in SCR_{sqr} change. *Note: change in expectations was operationalized as the difference between post-OBL and baseline expectations ratings in placebo trials; change in* SCR_{sqr} *was operationalized as the difference between post-OBL and baseline* SCR_{sqr} *in placebo trials; total, direct and indirect effects are not significant; the controlled variable is represented by the dotted lines; total effect is presented in brackets;* R^2 *for the model =0.24, F(4,83)=6.62, p<0.001; **p<0.01, ***p<0.001; ns – not significant.*

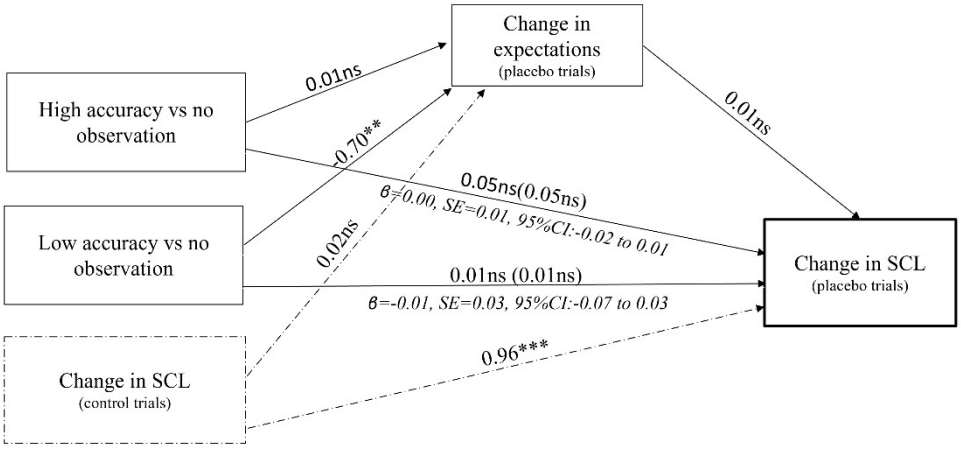

Figure S4. Results of analysis testing the mediating role of change in expectations in between-groups differences in SCL_{sqr} change. *Note: change in expectations was operationalized as the difference between post-OBL and baseline expectation ratings in placebo trials;change in SCL_{sqr} was operationalized as the difference between post-OBL and baseline SCL_{sqr} in placebo trials; total, direct and indirect effects are not significant; the controlled variable is represented by the dotted lines; total effect is presented in brackets; R^2 for the model =0.91, F(4,88)=226.40, p<0.001; **p<0.01, ***p<0.001; ns – not significant.*

Supplementary Tables

Table S1. Correlations between response to a placebo and individual characteristics within merged experimental groups

Individual characteristics	r (Pearson)	p
LOT-R Optimism	-0.03	0.84
SSIE/EQ short Total	0.02	0.90
SWE/IRI Empathic Concern	0.09	0.50
SWE/IRI Personal Distress	-002	0.86
SWE/IRI Perspective Taking	-0.05	0.70
SPSRQ Sensitivity to Rewards	-0.01	0.95
SPSRQ Sensitivity to Punishment	-0.04	0.76
FPQ - Total	-0.002	0.99
FPQ - Severe Pain	-0.03	0.81

FPQ - Minor Pain	0.01	0.95
FPQ - Medical Pain	-0.003	0.98
BAIAS - Body Ignorance	0.12	0.35
BAIAS - Body Awareness	-0.02	0.88
BAIAS - Body Attention	0.18	0.16

Note: N=59 for FPQ-Total, N=61 for FPQ – Severe Pain, N=63 for FPQ Medical Pain, N=64 for other correlations. Placebo effect was operationalized as the difference in pain ratings between trials (placebo minus control) in post-OBL in OBL groups; LOT-R – Life Orientation Test-Revised, EQ-Short -Empathy Quotient , IRI – Interpersonal Reactivity Index, FPQ – Fear of Pain Questionnaire, SPSRQ – Sensitivity to Punishment and Reward Questionnaire, BAIAS – Body Attention, Ignorance and Awareness Scale

Table S2. The effect of accuracy in pain reporting on the response to a placebo

Variable	Coefficient	SE	t	p	95% CI	
					Lower	Upper
Constant	-0.12	0.1	-0.82	0.42	-0.43	0.18

		5				
Group	-0.11	0.17	-0.67	0.50	-0.44	0.22
CVR	1.44	0.95	1.52	0.13	-0.44	3.32
CVR x group	-0.22	1.10	-0.20	0.84	-2.41	1.97
Change in pain (control trials)	0.56	0.11	5.16	<0.001	0.34	0.77

Note: CVR – average coefficient of variation for pain ratings. R^2 for the model: 0.27, $F(4,91)=8.18$, $p<0.001$, $\Delta R^2 = 0.00$, $F(1,91)=0.04$, $p=0.84$

Table S3. Control questions – between-group differences

Question	F value	p	ηp2
How accurate was the model when estimating pain?#	F(1,62)=324.64	<0.001	0.84
How accurate was the participant when estimating pain?	F(1,62)=46.80	<0.001	0.43
How did the ointment affect the model's pain?	F(1,62)=0.99	0.32	0.02
How did the ointment affect the participant's pain?	F(1,94)=1.12	0.29	0.01
How much pain did the model experience on average with the ointment?	F(1,62)=2.73	0.10	0.04
How much pain did the model experience on average without the ointment?	F(1,62)=1.77	0.19	0.03
How much pain did the participant experience on average with the ointment?	F(1,94)=0.41	0.53	0.00
How much pain did the participant experience on average without the ointment?	F(1,94)=5.57	0.02	0.06
How did the model's pain rating affect the participant's pain perception?	F(1,62)=0.51	0.48	0.01
How much did the participants try to adjust their pain ratings to the model's ratings?	F(1,62)=0.13	0.29	0.00

Note#: Since the correlation between the two measures of model's accuracy used in the current study was very high ($r=0.95$, $p<0.001$), we used only one of the measures (how accurately the model estimated the difference between the stimuli) in these comparisons.