Socially transmitted placebo effects

Pin-Hao A. Chen, Jin Hyun Cheong, Eshin Jolly, Hirsh Elhence, Tor D. Wager, & Luke J. Chang[*]

Department of Psychological and Brain Sciences
Dartmouth College
Hanover, NH, USA

*Corresponding author (luke.j.chang@dartmouth.edu)

Word count: 4860
Key words: placebo, interpersonal expectancy, social interactions

Abstract

Medical treatments typically occur in the context of a social interaction between healthcare providers and patients. Although, decades of research have demonstrated that patients' expectations can dramatically impact treatment outcomes, less is known about the influence of providers' expectations. Here, we systematically manipulated providers' expectations in a simulated clinical interaction involving administration of thermal pain and found that patients' subjective experiences of pain were directly modulated by providers' expectations of treatment success reflected in the patients' subjective ratings, skin conductance responses, and facial expression behaviors. The belief manipulation also impacted patients' perceptions of providers' empathy during the pain procedure and manifested as subtle changes in providers' face expression behaviors during the clinical interaction. Importantly, these findings replicated in two additional independent samples. Together, our results provide evidence of a socially transmitted placebo effect, highlighting the importance of how healthcare providers' behavior and cognitive mindsets can impact clinical interactions.

Introduction

The scientific foundation of medical treatments is based on the notion that alleviating patient symptoms requires treating the underlying biological disease process. However, it has been known since at least the 1930s that contextual and psychological factors such as expectations [1,2], and the characteristics of clinicians themselves [3–5], can profoundly impact symptom relief. The desire to minimize the impact of patients' expectations as well as clinicians' expectations that may be communicated to patients, has led to the adoption of the double-blind randomized clinical trial as the gold standard for testing drug treatments. In fact, reducing expectancy and other 'non-specific' treatment effects is a critical development in modern medicine and has accelerated the rapid advance of medical treatments in the 20th century [6,7]. Studies of placebo effects have demonstrated that manipulations of the interpersonal [8–14] and physical treatment context can, in some cases, produce substantial effects on symptoms and behavior [15–17] and associated brain processes [16,18–22]. Improvement in placebo groups in randomized clinical trials can rival even the most technologically advanced treatments for neuropsychiatric disorders, such as dopaminergic gene therapy for Parkinson's [23] and deep brain stimulation for depression [15,16]. Thus, an important part of standard 'open-label' clinical treatment includes both patients' and clinicians' expectations. However, considerably less is known about how clinicians' expectations are transmitted to patients and might ultimately impact their clinical outcomes.

Healthcare-providers' expectations of their ability to help a patient are particularly important to consider, as clinical trials are rarely truly double-blind. In psychotherapy trials, a provider cannot be blind to the treatment they are administering and therapists may favor a specific treatment, which can account for an estimated 69% of variance in treatment outcomes [24]. In medical clinical trials, there are often unblinding effects caused by the experimenters or even by treatment side effects [25,26], which can lead both patients and providers to accurately identify the treatment conditions to which the patients belong [5,27–29]. Thus, even within the context of randomized clinical trials, providers' expectations are likely contributing to successful treatment outcomes [30,31].

These effects may share similar mechanisms with interpersonal expectancy effects studied across areas of psychology, including psychotherapy, education, and beyond [32–34]. Specifically manipulating experimenters' expectations about treatment outcomes can have profound effects across a variety of experimental contexts. For example, rodents have been demonstrated to complete a maze significantly faster if the experimenter believes they have been bred to be more intelligent [35]. Similarly, schoolchildren will perform better on a standardized test at the end of the year if their teachers are led to believe that they are "growth spurters" [36]. Furthermore, manipulating experimenters' expectations about a psychological effect can lead to a self-fulfilling prophecy [37] even when the dependent measures are completely automated and outside the experimenter's control [38]. These interpersonal-expectancy results are remarkably robust and have a consistently large effect size (Cohen's D=.70) across hundreds of studies [32,33].

Despite the robustness of these interpersonal expectancy effects, there has been surprisingly little work demonstrating a causal link between providers' expectations and patients' treatment outcomes. Psychotherapy research has found modest correlations ($r^2 \sim .08$) between providers' expectations and treatment outcomes [39–41]. Only one study, to our knowledge, has attempted to demonstrate that provider's expectations may be transmitted to patients in clinical contexts using

a double-blind design [42]. This study specifically examined self-reported pain following dental surgery in two groups of patients that were randomly assigned to receive a placebo. All patients believed that they would randomly receive a drug that would decrease their pain (fentanyl), increase their pain (naloxone), or have no effect on their pain (placebo). In one placebo group (n=18), the treating providers believed the patients had a 33% chance of being randomized to an analgesic, while providers treating the other placebo group (n=8) believed there was a 0% chance of being randomized to an analgesic. Interestingly, only the placebo group, which the providers believed they could receive fentanyl reported a decrease in pain 60 minutes post treatment. This provocative study suggests that providers' expectations of treatment effectiveness may impact patients' treatment outcomes.

In the present study, we systematically test for an interpersonal expectancy effect in a simulated clinical interaction involving administration of thermal pain. In a single-blind design, we examined the impact of the provider's expectations of the analgesic effect of two different treatment creams on the patient's pain experience. Across three studies, participants (N=194) were randomly assigned to play the role of either a 'doctor' or 'patient'. Doctors were told that they would be administering "Thermedol", a TRP-channel blocker with analgesic effects for thermal pain, and an inert control cream to patients as a treatment to mitigate the effects of noxious heat stimulation. In actuality, both creams were identical petroleum-based jelly with no analgesic effects. In addition to being instructed about the effectiveness of each cream, doctors underwent a placebo conditioning protocol, in which each cream was paired with different temperatures of thermal stimulation [43,44]. In the subsequent interaction phase, doctors administered each cream to the volar surface of patients' forearms and then applied a 47C thermal stimulation. We examined whether the doctors' beliefs impacted: (a) patients' subjective reports of pain experience, (b) patient's autonomic arousal measured through their skin conductance response (SCR), and (c) patients' pain behavior communicated via facial expressions. Importantly, through this systematic manipulation, we explored an unknown territory in clinical settings - how doctors transmit their beliefs to patients - providing an empirical demonstration of a socially transmitted placebo effect.

Results

Doctor Conditioning Phase

In Study 1, we recruited 48 participants (52.08% female) comprising a total of 24 dyads and randomly assigned one person to the role of doctor and the other the role of patient. In the Doctor Conditioning phase, we manipulated the doctor to believe that the "Thermedol" cream was more effective in reducing thermal pain compared to the Control cream. The doctor was asked to experience the effectiveness of each cream using a conditioning paradigm such that Thermedol was paired with a lower temperature (43 Celsius) compared to the Control treatment (47 Celsius) (Figure 1(A-1)). Overall, doctors reported experiencing less thermal pain when Thermedol was applied compared to the Control treatment, b^1 = -22.63, *SE* = 2.91, *t*(21.95) = -7.78, *p* < .001, CI[-28.45, -16.81] (Figure 1(A-3), Extended Data Figure 1), which corresponded to increased beliefs in the effectiveness of the Thermedol treatment, *b* = 24.60, *SE* = 5.08, *t*(23.00) = 4.84, *p* < .001, CI[34.76 ,14.44] (Figure 1(A-2), Extended Data Figure 1). Beta values reflect the difference in

rating scores when doctors received the Thermedol treatment compared to the Control treatment (i.e., b^1 = 22.63 indicates the difference in scores between conditions on a 100 point scale), we assumed a normal distribution of data and used two-tailed hypothesis tests.
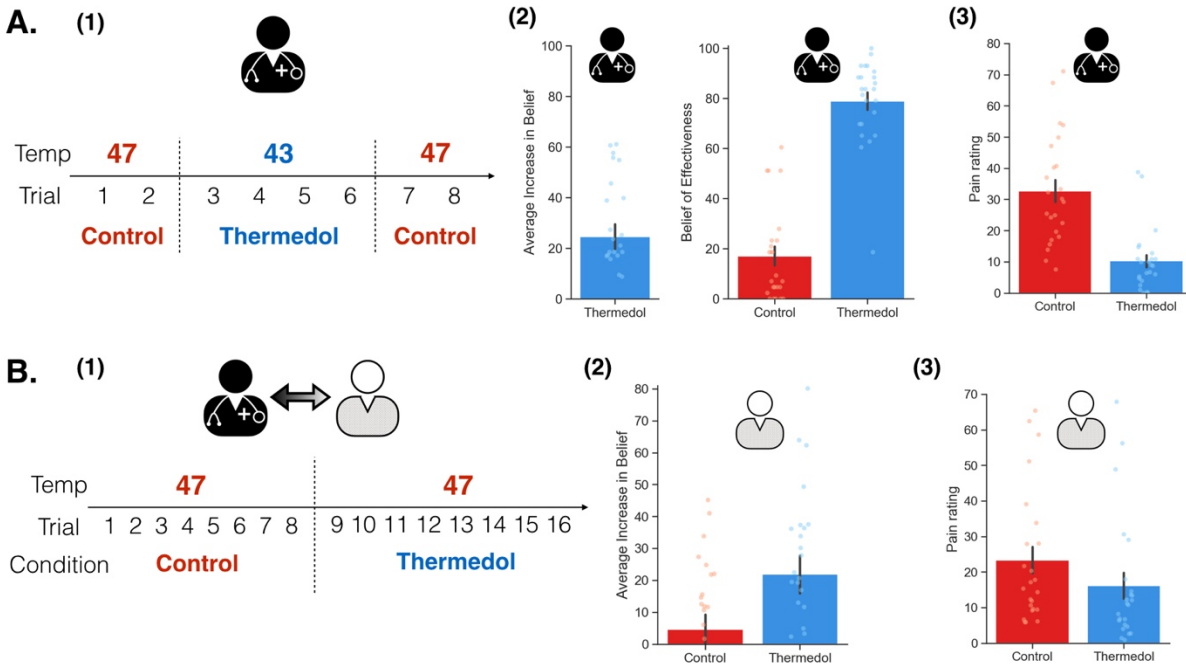

Figure 1. Experimental design and subjective reports of pain and beliefs of effectiveness in Study 1. (A) During the Doctor Conditioning phase, (1) the doctor was manipulated to believe that the Thermedol treatment (blue) reduced thermal pain more than the Control treatment (red) by pairing a lower temperature with the Thermedol treatment over 8 trials; (2) conditioning was successful as the doctor reported an increased belief in the efficacy of Thermedol, b = 24.60, SE = 5.08, $t(23.00)$ = 4.84, $p < .001$, CI[34.76 ,14.44], as well as maintained a stronger belief to the Thermedol in the next phase, b = 61.92, SE = 5.60, $t(26.54)$ = 11.05, $p < .001$, CI[73.12, 50.72]; and (3) the doctor experienced less thermal pain with the Thermedol than the Control treatment, b = -22.63, SE = 2.91, $t(21.95)$ = -7.78, $p < .001$, CI[-28.45, -16.81]. **(B)** During the Doctor-Patient Interaction phase, (1) the patient interacted with the doctor, but the patient received the same intensity of thermal stimulation pain across both treatments; (2) in line with the doctor's reports, the patient reported a stronger belief in the efficacy of Thermedol, F(1, 23.00) = 5.63, p = .03; and (3) the patient experienced less thermal pain with the Thermedol than the Control treatment, b = -7.30, SE = 1.53, $t(22.00)$ = -4.78, $p < .001$, CI[-10.36, -4.24]. All panels include data from 24 dyads. Error bars represent standard error of mean (S.E.M.).

Doctor-Patient Interaction Phase

Next, in the doctor-patient interaction phase, we examined whether the doctors' beliefs about the treatment being administered impacted the patient's subjective experience of pain. Doctors were instructed that this experiment followed a single-blind procedure, and that they were not permitted

to reveal differences between these two treatments when interacting with the patients. Importantly, doctors' beliefs formed in the Conditioning phase were maintained in the Interaction phase before they administered the treatments, b = 61.92, SE = 5.60, $t(26.54)$ = 11.05, $p < .001$, CI[73.12, 50.72] (Figure 1(A-2) & Extended Data Figure 2(C), Supplementary Table 1).

Across both treatments, patients received the same temperatures of pain stimulation (i.e. 47 Celsius) and were instructed to rate their belief of the effectiveness of each treatment and their overall and continuous pain ratings. Following previous placebo-conditioning studies that have demonstrated counterbalancing effects indicating the need for a reference experience prior to receiving a placebo [43], patients always received the Control before the Thermedol treatment (Figure 1, B-1)). We found that despite receiving the same level of thermal stimulation, patients reported a significant increase in beliefs of effectiveness in the Thermedol compared to Control treatment, $F(1, 23.00)$ = 5.63, p = .03 (Figure 1(B-2), Supplementary Table 1), and consistent with these beliefs, patients reported experiencing less pain in the Thermedol treatment compared to Control, b = -7.30, SE = 1.53, $t(22.00)$ = -4.78, $p < .001$, CI[-10.36, -4.24] (Figure 1(B-3), Supplementary Table 1). This finding was also observed in the maximal pain reported during their continuous pain ratings, b = -11.52, SE = 2.05, $t(22.00)$ = -5.63, $p < .001$, CI[-15.62, -7.42] (Extended Data Figure 2(B), Supplementary Table 1).

Objective Measures of Pain Experience

In order to rule out the possibility that the observed socially transmitted placebo effect was not simply a reporting bias akin to a demand characteristic [45], we examined the effect of manipulating doctors' beliefs on two objective measures of pain experience: autonomic arousal (SCR), and behavioral displays of pain (facial expressions). Consistent with their subjective experience of pain, patients displayed a lower SCR response for Thermedol compared to the Control treatment, b = -1.67, SE = 0.52, $t(20.14)$ = -3.20, p = .004, CI[-2.71, -0.63] (Figure 2(A), Supplementary Table 1).

Both patients' and doctors' facial expressions were recorded at 120hz using custom headsets [46]. We used a pre-trained convolutional neural network model to extract 20 facial action units plus three estimated head rotations (pitch, yaw, and roll) for each frame of the video during the experiment [47,48]. Facial action units are a standardized system to describe the intensity of facial muscle movements [49]. We then trained a pain expression model (PE model) predicting the doctor's overall pain ratings during each trial of the conditioning phase using a linear ridge regression with three features of each facial action unit (i.e., the probability of maximal expression (max), probability of minimal expression (min), and time to maximal expression (tmax)). Using leave-one-subject-out (LOSO) cross-validation, this model achieved an average correlation between predicted and actual pain ratings of r = .41, r^2 = .17, SD = .33, p = .003, determined by permuting training labels (Extended Data Figure 3(B); see Methods for details). We projected the model weights from the max value onto a two dimensional representation of facial landmarks. This visualization suggests that the model is detecting facial behaviors previously associated with pain expression as features related to brow lowerer (AU4), lip corner puller (AU12), wrinkling of

nose (AU9), and raising upper lip (AU10) predicted greater pain ratings (Figure 2(B-1), Extended Data Figure 3(A), Supplementary Table 2) [50,51].
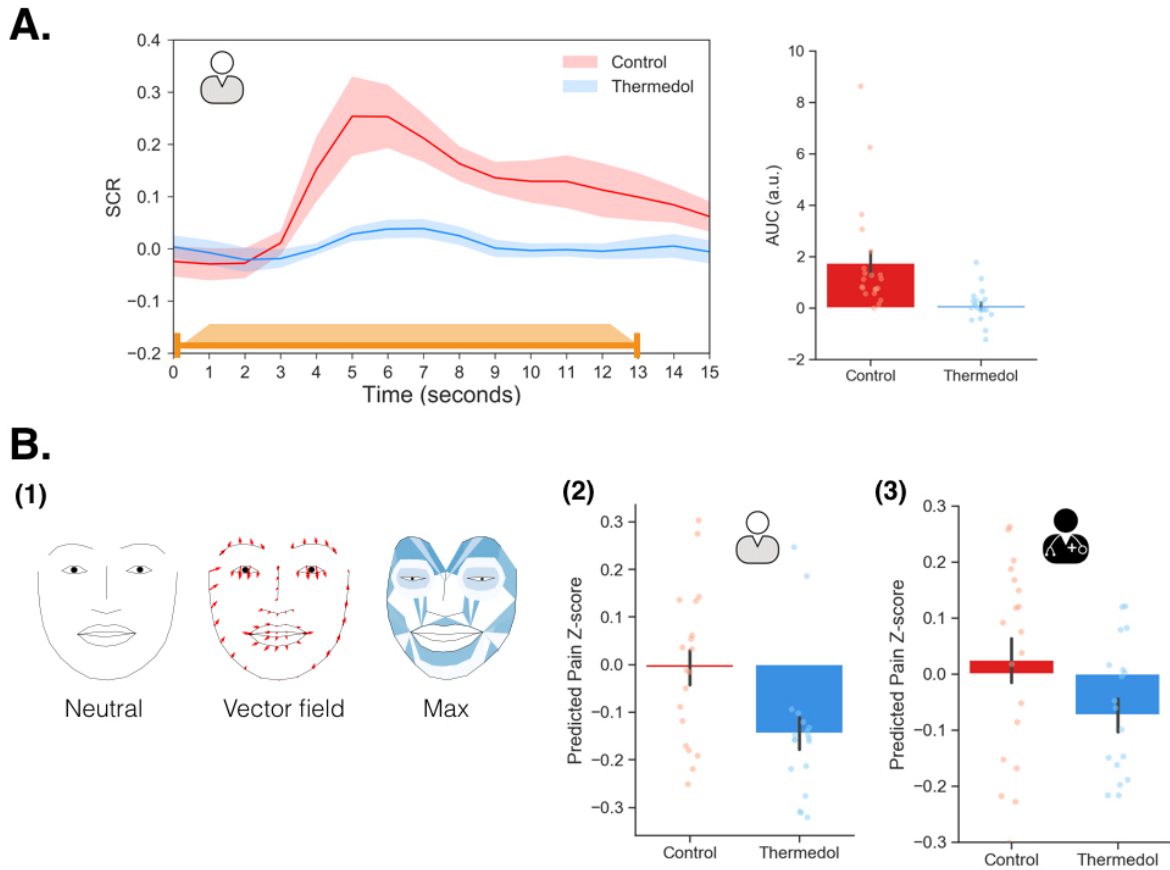

Figure 2. Objective measures of pain experience in study 1. (A) The orange line indicates the duration of pain stimulation (1.5 seconds ramp-up, 10 seconds at peak temperature, and 1.5 seconds ramp-down). The patient showed lower SCR to the Thermedol (blue) than the Control treatment (red), b = -1.67, SE = 0.52, $t(20.14)$ = -3.20, p = .004, CI[-2.71, -0.63]. **(B)** (1) A visualization of the pain expression (PE) model; the left face demonstrated a neutral facial expression; the middle face depicts the 2D deformation of each facial landmark from the neutral expression to the PE model. The red arrows indicate the vector field of how each landmark morphed from neutral to the pain expression; the right face depicts the overall deformation of the face along with how much the maximum value of each AU contributed to predicting pain in the PE model (darker color indicates higher parameter estimates) (2) The patient expressed reduced pain facial behaviors under the Thermedol than the Control treatment, b = -0.14, SE = 0.05, $t(17.02)$ = -2.86, p = .01, CI[-0.24, -0.04]. (3) In line with the patient, the doctor also expressed less pain facial behaviors to the Thermedol than the control treatment, b = -0.10, SE = 0.03, $t(17.78)$ = -2.43, p = .03, CI[-0.16, -0.04]. Panel A includes data from 21 patients, and panel B includes data from 19 dyads. Error bars represent S.E.M.

We next tested whether facial expression was a plausible mechanism for the communication of beliefs from doctors to patients. The PE model, trained on *doctors'* pain facial expressions, was also able to reliably predict *patients'* reported experiences of pain. Specifically, model predictions significantly correlated with patients' overall pain ratings, r = .24, SD = .31, p = .003, determined by a permutation test (Extended Data Figure 3(C); see Methods for details), validating the reliability and generalizability of the PE model. Importantly, predicted pain ratings in the facial expression-based model was significantly lower in the Thermedol treatment than in the Control treatment, b = -0.14, SE = 0.05, t(17.02) = -2.86, p = .01, CI[-0.24, -0.04] (Figure 2(B-2); Supplementary Table 1). Together, these results provide converging evidence across self-reported pain, autonomic arousal, and facial behaviors that patients were experiencing less physical pain when doctors believed they were delivering an efficacious treatment.

Doctor's display different facial expressions between Thermedol and Control treatments

Next, we examined how Doctors might be transferring their beliefs about the treatments to patients via doctor's facial expressions. One possibility is that the doctors were displaying different pain facial expressions when delivering the Thermedol treatment compared to the control because they believed that Thermedol was the only effective treatment. To test this hypothesis, we applied the PE model to predict doctors' pain facial behaviors in the Doctor-Patient interaction phase. Interestingly, we observed a similar effect to the patients, doctors expressed less pain facial expression behavior while patients were receiving stimulation with the Thermedol than with the Control treatment, b = -0.10, SE = 0.03, t(17.78) = -2.43, p = .03, CI[-0.16, -0.04] (Figure 2(B-3), Supplementary Table 1). In addition, patients appeared to be sensitive to this behavioral change in the doctors, and reported finding the doctors more empathetic in the Thermedol treatment compared to the Control, b = 9.61, SE = 3.36, t(33.67) = 2.86, p = .007, CI[16.33, 2.89] (Extended Data Figure 2(D), Supplementary Table 1).

To assess if the doctor's facial expressions served as the mechanism of the belief transmission, we conducted two additional mediation analyses. First, we examined if the transmission of beliefs between the doctor and patient was mediated by the doctor's facial expressions. We found that our manipulation of doctors' beliefs (condition) was highly related to patients' subjective pain ratings, b = -7.03, SE = 1.99, t(18.00) = -3.54, p = .002, CI[-11.01, -3.05], but this relationship did not appear to be mediated by the predicted pain level from doctors' facial expressions, b = -8.11, SE = 2.31, t(18.78) = -3.52, p = .002, CI[-12.73, -3.49] (with the mediator in the model). Second, we explored whether changes in the doctors' facial expressions could account for the patients' empathy effects. We found that our manipulation of doctors' beliefs (condition) was highly related to patients' perceptions of doctors' empathy, b = 12.75, SE = 3.86, t(18.00) = 3.31, p = .004, CI[20.47, 5.03], but this relationship did not appear to be mediated by the predicted pain level from doctors' facial expressions, b = 13.34, SE = 4.14, t(18.19) = 3.02, p = .007, CI[21.62, 5.06] (with the meditator in the model). However, the predicted pain level from doctors' facial expressions was marginally related to patients' perceptions of doctors' empathy, b = -37.50, SE = 26.13, t(33.78) = -1.44, p = .16, CI[14.76, -89.76]. These results suggest that though the doctors are behaving differently between the Thermedol and Control conditions, the amount of predicted

pain in the doctors' facial expressions did not directly lead to the patients' reduced pain experience.

Study 2

In Study 2 (n=86 participants; 62.79% female), we attempted to address additional questions about the socially transmitted placebo effect. First, Study 1 had a modest sample size and we wanted to ensure the robustness of our effects by replicating the effect in an independent sample. Second, we wanted to rule out the possibility that our socially transmitted placebo effects were not simply due to participants habituating to the pain stimulation. For example, participants always experienced less pain in the treatment delivered in the second order. Thus, in Study 2, we counterbalanced the treatment order across dyads, where some dyads received the Control treatment first (original order) whereas the others received the Thermedol first (reverse order). Third, we additionally examined the effect of temperature intensity to test the consistency of our findings across two different temperatures (46 and 47 Celsius) and also examined the impact of experimenter using two separate experimenters.

Overall, we were able to successfully replicate the results from Study 1 (Figure 3(A-1)). In the conditioning phase, doctors reported increased beliefs of effectiveness to the Thermedol, b = 16.49, SE = 4.98, $t(42.00)$ = 3.35, p = .002, CI[26.45, 6.53] (Figure 3(A-2), Extended Data Figure 4) and experienced less pain, b = -22.93, SE = 2.28, $t(40.05)$ = -10.07, p < .001, CI[-27.49, -18.37] (Figure 3(A-3), Extended Data Figure 4) in the Thermedol compared to the Control treatment condition. In the Doctor-Patient interaction phase, we found a significant Time x Condition x Order interaction, $F(1, 81.58)$ = 9.99, p = .002 (Supplementary Table 3), in patients' beliefs of effectiveness. Patients reported higher beliefs of effectiveness for the Thermedol compared to the Control treatment after pain stimulation only when treatments were delivered in the original order (i.e. Control followed by Thermedol), b = 24.58, SE = 6.13, $t(82.66)$ = 4.01, p = .003, CI[36.84, 12.32] (Figure 3(B-2)), but not in the reverse order, b = -9.00, SE = 6.57, $t(82.66)$ = -1.37, p = .87, CI[-22.14, 4.14] (Figure 3(C-2)). We also observed a significant Condition x Order interaction for the patient's subjective pain experience, $F(1, 38.10)$ = 12.34, p = .001 (Supplementary Table 3), with patients reporting experiencing less pain for the Thermedol than the Control treatment only in the original order, b = -7.35, SE = 1.84, $t(38.46)$ = -4.00, p = .002, CI[-11.03, -3.67] (Figure 3(B-3)), but not in the reverse order, b = 2.25, SE = 2.05, $t(38.47)$ = 1.10, p = 0.69, CI[6.35, -1.85] (Figure 3(C-3)). Similarly, patients also showed a significant Condition x Order interaction, $F(1, 398.55)$ = 11.00, p < .001 (Supplementary Table 3) for SCR, exhibiting lower SCR for Thermedol than the Control treatment, but only in the original order, b = -0.86, SE = 0.37, $t(379.35)$ = -2.81, p = .03, CI[-1.60, -0.12] (Extended Data Figure 5(A-2)), and not in the reverse order, b = 0.53, SE = 0.29, $t(375.93)$ = 1.85, p = .25, CI[1.11, -0.05] (Extended Data Figure 5(B-2)). In addition, patients reported finding the doctors more empathetic in the Thermedol treatment compared to the Control, b = 15.15, SE = 5.67, $t(84.76)$ = 2.67, p = .009, CI[26.49, 3.81], and this finding was not affected by the administration order of the two treatments, Condition x Order interaction, $F(1,42.48)$ = 0.43, p = .51 (Supplementary Table 3). Finally, we also observed a significant main effect of temperature intensity on subjective pain experience, $F(1, 39.12)$ =

8.00, p = .007, but no main effect of experimenter on pain intensity, nor any significant interactions with the conditions of interest (Supplementary Table 3).
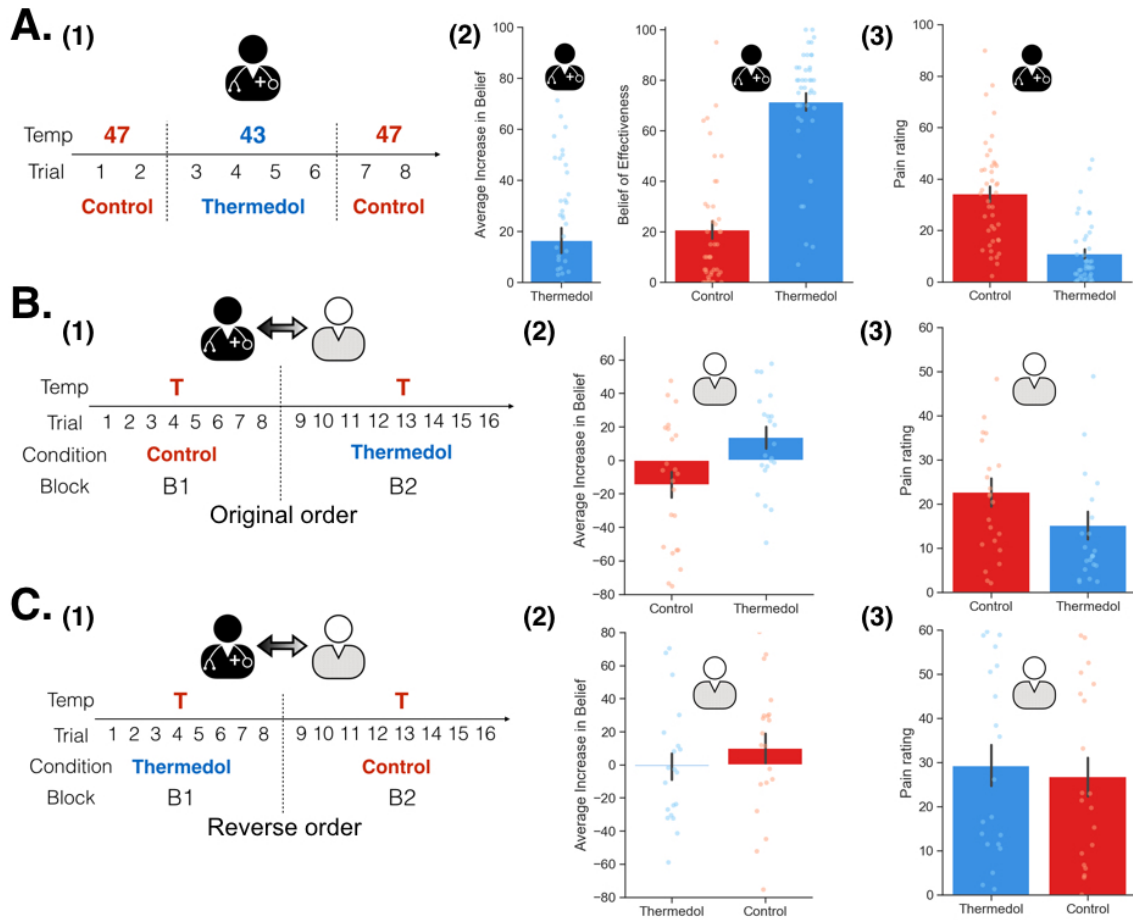

Figure 3. Experimental design and subjective reports of pain and beliefs of effectiveness in Study 2. (A) During Doctor Conditioning phase, (1) Doctors were manipulated to believe that the Thermedol treatment (blue) reduced more thermal pain than the Control treatment (red) by pairing a lower temperature with the Thermedol treatment. (2) The doctors reported increased beliefs in the effectiveness of the Thermedol treatment, b = 16.49, SE = 4.98, $t(42.00)$ = 3.35, p = .002, CI[26.45, 6.53], which was maintained in the next phase, and (3) the doctors experienced less thermal pain with the Thermedol than the Control treatment, b = -22.93, SE = 2.28, $t(40.05)$ = -10.07, p < .001, CI[-27.49, -18.37]. **(B)** During the Doctor-Patient Interaction phase, (1) half of the patients received treatments in the original order (the same order as Study 1) with the same intensity of pain. (2) The patients reported having higher increases in beliefs of effectiveness, b = 24.58, SE = 6.13, $t(82.66)$ = 4.01, p = .003, CI[36.84, 12.32], and (3) experiencing less thermal pain with the Thermedol than the Control treatment, b = -7.35, SE = 1.84, $t(38.46)$ = -4.00, p = .002, CI[-11.03, -3.67]. **(C)** (1) Half of the patients received treatments in the reverse order (Thermedol first and Control second). (2) The patients reported having no difference between the two treatments in beliefs of effectiveness, b = -9.00, SE = 6.57, $t(82.66)$ = -1.37, p = .87, CI[-22.14, 4.14], and in (3) subjective experience of thermal pain, b = 2.25, SE = 2.05, $t(38.47)$ = 1.10,

p = 0.69, CI[6.35, -1.85]. All panels include data from 43 dyads across both orders. Error bars represent S.E.M.

In summary, these results fully replicated all effects observed in Study 1 in an independent sample and provide additional support that the effect is unlikely to be a result of specific participants, or a specific experimenter. Alhough we did not observe evidence of a habituation effect, it is important to note that we did find that the temporal ordering of the conditions significantly impacted the results, which is consistent with previous placebo conditioning studies [43].

Study 3

In Study 3 (n = 60 participants; 65.00% female), we attempted to address central remaining questions from Study 1 and 2 about the socially transmitted placebo effect. First, although findings from Study 2 suggested that this effect was unlikely to be explained by habituation in a between-subject design, a within subject design would provide an even stronger test. Thus, in Study 3, we employed a within-subject ABBA design (administration order: Control - Thermedol - Thermedol - Control) during the Doctor-Patient interaction phase. Second, to confirm that our findings were not impacted by the presence of cameras, we removed head-mounted cameras in this study. Third, since we did not control for the expectations of experimenters in the first two studies, in Study 3, experimenters were completely blind to the experimental conditions during the Doctor-Patient interaction phase. Finally, we sought to examine whether our findings could be extended using a higher thermal stimulation temperature (i.e., 48 Celsius).

We were again able to successfully replicate the results from Study 1 and Study 2 (Figure 4). In the conditioning phase, doctors reported increased beliefs of effectiveness to the Thermedol, *b* = 25.89, *SE* = 3.95, *t*(29) = 6.55, *p* < .001, CI[33.78, 17.98] (Figure 4(A-2), Extended Data Figure 6) and experienced less pain, *b* = -31.92, *SE* = 2.89, *t*(30.02) = -11.02, *p* < .001, CI[-37.70, -26.14] (Figure 4(A-3), Extended Data Figure 6) in the Thermedol compared to the Control treatment condition. In the Doctor-Patient interaction phase, patients reported a significant increase in beliefs of effectiveness to the Thermedol compared to the Control treatment, Time x Condition interaction, *F*(1, 178.00) = 4.98, *p* = .03 (Figure 4(B-2), Supplementary Table 4), and in line with their beliefs, patients also reported experiencing less pain in the Thermedol relative to the Control treatment, *b* = -3.70, *SE* = 1.53, *t*(27.30) = -2.42, *p* = .02, CI[-6.76, -0.64] (Figure 4(B-3), Supplementary Table 4). Similarly, patients also showed lower SCR for the Thermedol than the Control treatment, *b* = -1.37, *SE* = 0.67, *t*(285.60) = -2.04, *p* = .04, CI[-2.71, -0.03] (Extended Data Figure 7(B), Supplementary Table 4). Finally, patients also reported finding the doctors more empathetic in the Thermedol compared to the Control treatment, *b* = 6.88, *SE* = 2.89, *t*(96.03) = 2.38, *p* = .02, CI[12.66, 1.10] (Supplementary Table 4).
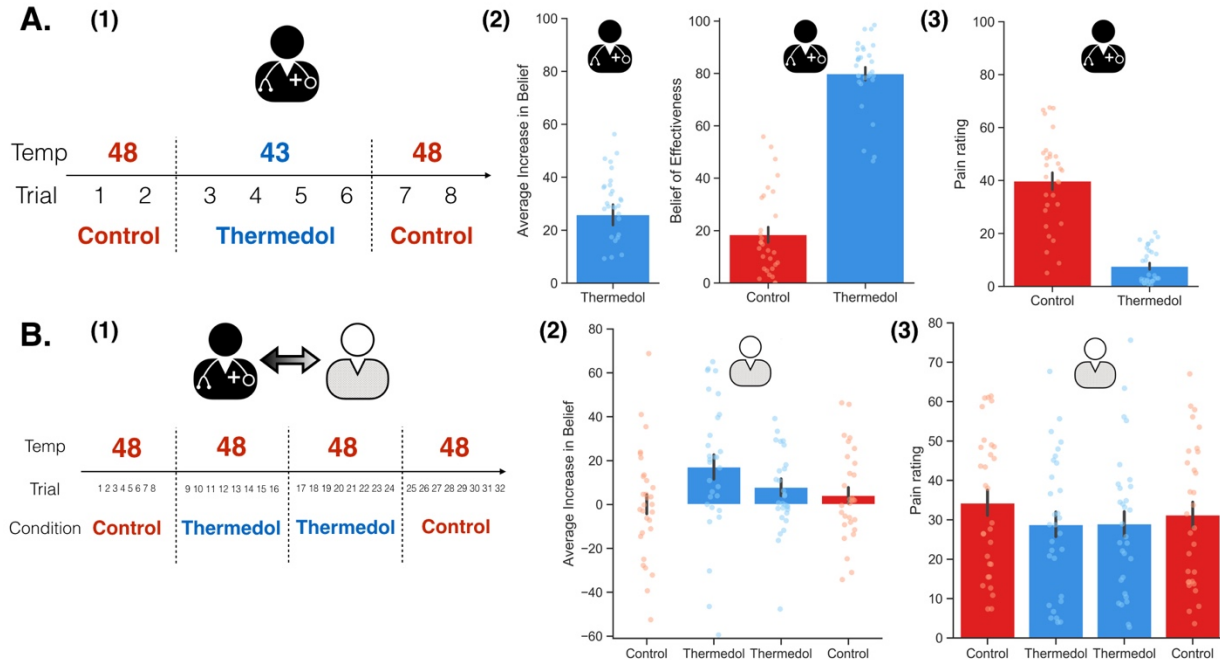
Figure 4. Experimental design and subjective reports of pain and beliefs of effectiveness in Study 3. (A) During the Doctor Conditioning phase, (1) the doctors were manipulated to believe that the Thermedol treatment (blue) reduced more thermal pain than the Control treatment (red) by pairing a lower temperature with the Thermedol treatment, (2) the doctor reported having an increased beliefs of effectiveness in the Thermedol, $b = 25.89$, $SE = 3.95$, $t(29) = 6.55$, $p < .001$, CI[33.78, 17.98], as well as maintained a stronger belief to the Thermedol in the next phase, $b = 62.57$, $SE = 3.37$, $t(37.18) = 18.57$, $p < .001$, CI[69.31, 55.83], and (3) the doctor experienced less thermal pain with the Thermedol than the Control treatment, $b = -31.92$, $SE = 2.89$, $t(30.02) = -11.02$, $p < .001$, CI[-37.70, -26.14]. **(B)** During the Doctor-Patient Interaction phase, (1) the patient interacted with the doctor, but the patient received the same intensity of thermal stimulation across all trials in both treatments, (2) consistent with doctor's beliefs, the patient showed an increase in beliefs of effectiveness of Thermedol, $F(1, 178.00) = 4.98$, $p = .03$, and (3) the patient experienced less pain with the Thermedol than the Control treatment, $b = -3.70$, $SE = 1.53$, $t(27.30) = -2.42$, $p = .02$, CI[-6.76, -0.64]. All panels include data from 30 dyads. Error bars represent S.E.M.

Overall, the results in Study 3 successfully replicated all effects from Study 1 and Study 2 in another independent sample. Importantly, we demonstrated that the effects could not be explained by central or peripheral habituation using a within-subject ABBA design. We also showed that the effects could be generalized to contexts without the presence of cameras, which are more similar to actual clinical contexts and suggests that the effect was not enhanced by increased awareness that participants were being video monitored. Lastly, we controlled for the expectations of the experimenters, demonstrating that the effects were purely transmitted from doctors to patients. These findings provide strong evidence that the socially transmitted placebo effect is not due to habituation, presence of cameras, and beliefs of experimenters.

Discussion

In this study, we systematically examined the impact of doctors' beliefs about treatment effectiveness on patients' pain experience in a single-blind simulated clinical interaction across three independent studies. We successfully induced doctors' beliefs about the efficacy of the Thermedol treatment via instruction and a conditioning protocol. These beliefs were implicitly transmitted to the patients during a brief single-blind treatment. After the clinical interaction with doctors, patients not only showed the same beliefs of treatment effectiveness, but also subjectively perceived less pain as a result of the treatment despite there being no real difference between treatment and control conditions. Importantly, this did not appear to be a reporting bias as this pattern of results was also observed using objective pain measures, such as psychophysiology (i.e. SCR) and facial expressions. These results appear to generalize beyond the specific doctors, patients, and experimenters as we successfully replicated our effects in an independent sample in Study 2. In Study 3, we further demonstrated that these results were not due to central or peripheral habituation, nor influenced by the presence of cameras. Importantly, by recruiting experimenters blind to the experimental conditions to run the interaction phase, we ensured that the effect was purely transmitted from the doctors to the patients. These findings provide a compelling demonstration that beliefs about a treatment's effectiveness induced in doctors can be implicitly transferred to patients via brief social interactions, which has important implications for how healthcare providers interact with patients.

Our findings provide strong evidence for the causal role of doctors' expectations of treatment effectiveness in patients' treatment outcomes, demonstrating the importance of interpersonal expectancy effects in clinical settings. Interpersonal expectancy effects have been observed across diverse domains, such as the effect of experimenters' expectations on the performance of rodents [35], the impact of experimenters' expectations on subjects' blood pressure [52,53], the influence of teachers' expectations on students' performance [36], and even the impact of judges' expectations on jurors' verdicts [54]. Although the robustness of interpersonal expectancy effects has been well established across hundreds of human and animal studies [32,33], there has been a paucity of research exploring this effect within clinical contexts [42]. Interestingly, a few studies have provided positive evidence that healthcare providers' or caregivers' expectations could potentially impact patients' treatment outcomes. For example, the mental status of nursing-home patients was shown to differ based on manipulating the expectations of caregivers; patients' mental status in the high expectancy group was better than those in the average expectancy group after weeks of treatments [55]. Similarly, caregivers' expectations further influenced patients' quality of life following cancer treatment [56]. Our study provides additional empirical support that interpersonal expectancies can substantially impact patients' treatment outcomes.

This work also provides additional empirical support for understanding the mechanisms underlying psychotherapeutic interventions. Meta-analyses of hundreds of psychotherapy studies have found that psychotherapy is an effective treatment compared to waitlist control, but no specific technique appears to consistently perform better than another [57]. While specific psychotherapeutic treatments have been estimated to only account for approximately 10% of treatment outcome variances, factors common to many treatments have been estimated to

explain approximately 70% of treatment outcome variance [58]. These so-called "common factors" include treatment mechanisms such as: patient-provider relationships, patients' expectations, and providers' expectations [59] and can be found in virtually all healing relationships [3,58,60]. Expectations that providers can help treat a malady [60] and a trusting therapeutic relationship [61] have been shown to be particularly effective in improving treatment outcomes by instilling hope [62], increasing treatment adherence [63], and even reducing malpractice lawsuits [64]. Moreover, treatment providers appear to have considerable variability in their overall treatment effectiveness even when delivering identical treatments, which provides additional support for the importance of expectations and relationships in treatment outcomes [65–67]. Our study demonstrates that common factors, such as provider expectations, can be studied empirically in a laboratory setting, and we hope that this will inspire many more rigorous investigations into these important mechanisms of change resulting from psychotherapeutic interventions.

How do healthcare providers implicitly transmit their beliefs to patients? We found evidence of subtle changes in doctors' facial expressions in response to the treatment based on their prior beliefs. Doctors appeared to convey more information relating to facial displays of pain when administering the Control treatment, which they believed was ineffective. We speculate that there are at least two possible mechanisms underlying the socially transmitted placebo effect. One possibility is that the doctor sends some nonverbal information that communicates their beliefs about which treatment is more likely to be the active treatment. Alternatively, the doctor might be more attentive to the patient, and develop a more empathetic connection which ultimately makes the patient feel better when they are receiving pain stimulations. The patient may then incorrectly attribute this analgesic effect to the Thermedol treatment. These proposed mechanisms are consistent with previous work emphasizing the importance of provider-patient alliances on patients' treatment outcomes [31,68–71]. For example, when patients perceived their providers as more empathetic during treatments, patients' severity and duration of the common cold decreased significantly compared to those who perceived their providers as less empathetic [72]. We believe that a more detailed understanding of the specific mechanisms of belief transmission should be explored in future work [73].

Although doctors' expectations were crucial in establishing the socially transmitted placebo effect, patients appeared to first require a reference experience before receiving the placebo treatment. Reference points have been demonstrated to play a critical role in sensory perception and subjective value. For example, early work on sensation and perception found an exponential relationship between the objective intensity of a stimulus and the perceived change in subjective perception referred to as the "just noticeable difference" [74,75]. Similar exponential effects have also been found in reference dependent models of utility such as prospect theory [76]. In clinical contexts, patients' symptoms (e.g. depressed mood) reflect their baseline state and serve as a reference experience, which can change based on successfully responding to a treatment often in the form of an exponential function. We speculate that in the absence of a pathological condition, healthy controls require a baseline reference experience before they can experience placebo induced analgesia. Indeed, several studies have demonstrated that healthy participants only exhibit a placebo effect when a control treatment is administered first [43,77,78]. These results informed our experimental design in Study 1 in both the Doctor Conditioning phase (i.e. high pain first and low

pain second) and the Doctor-Patient Interaction phase (i.e. control treatment first and placebo treatment second). Importantly, we extend this finding to socially transmitted placebos in Study 2. Individuals who received the original order showed robust placebo analgesia effect, whereas those who received the reverse order showed no such effect. We did not observe any evidence of habituation, but rather that the socially transmitted placebo first required a reference baseline pain experience.

One potential limitation of our study is that all participants were college students, not real healthcare providers or patients. Although the generalizability of our findings to clinical population may be limited, the fact that even participants with no medical expertise or training showed transmitted placebo effects through a brief social interaction in a highly controlled laboratory environment indicates that the influence of doctors' expectations on patients' outcomes will likely be even larger. In a real clinical context, more contextual cues, such as verbal suggestions and environmental cues, as well as doctors' and patients' prior experiences may amplify the transmission effect [79]. We look forward to future work that will expand on these findings in clinical contexts with real providers and patients.

In conclusion, we find converging evidence across three independent samples that providers' expectations about the efficacy of a treatment can substantially impact patients' treatment outcomes via implicit social cues. This finding has important implications for virtually all clinical interactions between patients and providers and highlights the importance of explicit training in bedside manner when delivering information and interventions. For example, some trainings focus on providers' psychological aspects, such as empathy [80–82], and others emphasize communication skills [83]. We believe that the tremendous resources invested in discovering novel treatments should be complemented by additional investment in understanding the mechanisms underlying one of the oldest and most powerful medical treatments - healers themselves.

Methods

Participants

A total of 194 participants (age range 18-28) were recruited from a large sample of undergraduate students from Introduction to Psychology and Introduction to Neuroscience classes at Dartmouth College and received course credit for their participation. Forty-eight participants (24 dyads, 52.08% female) were recruited in Study 1, 86 participants (43 dyads, 62.79% female) were recruited in Study 2 during a different term, and another 60 participants (30 dyads, 65.00% female) during a separate term. All participants gave informed consent in accordance with the guidelines set by the Committee for the Protection of Human Subjects at Dartmouth College.

Procedures

Study 1

Pre-experimental Instruction Phase

The experiment involved three phases: (1) the Pre-Experimental Instruction Phase, (2) the Doctor Conditioning Phase, and (3) the Doctor-Patient Interaction Phase. During the Pre-Experimental Instruction Phase, the experimenter told each dyad that since the goal of the experiment was to study social interaction in clinical settings, one participant was randomly assigned the role of 'doctor' and the other the role of 'patient'. The experimenter explained that participants would see two treatments during the experiment: one was Thermedol, a cream targeting skin pain receptors to reduce thermal pain (selectively blocks transient receptor potential (TRP) ion-channels to reduce nociceptive pain by disrupting heat sensing [84]), and the other one was an inert Vaseline cream, serving as the Control treatment. However, in actuality, both treatments were inert Vaseline cream. Participants were then instructed that the doctor would deliver these two treatments to the patient in an unknown order, and the patient needed to report their pain ratings while receiving thermal pain stimulation from the doctor. Lastly, the experimenter told the dyad that before the patient received treatments, the doctor would participate in a practice trial with the experimenter so that they could familiarize themselves with the experimental procedure. In fact, this practice trial served as the Doctor Conditioning Phase, making the doctor believe one treatment reduces more pain than the other one.

Doctor Conditioning Phase

At the beginning of the Doctor Conditioning Phase, the doctor was first told that they would experience an abbreviated version of the experimental procedure used in the Doctor-Patient interaction phase. The doctor was manipulated to believe that Thermedol more effectively reduced thermal pain compared to the control. The success of manipulating doctors' belief relied on two stages. First, the experimenter squeezed Thermedol from a seemingly authentic tube (i.e. with an ingredient list, a UPC code, manufacturer information) into one container and instructed the doctor to mix Vaseline cream with food coloring into a separate, but identical container. Each treatment was associated with a different color (i.e., blue, red) in order to help the doctors differentiate the two treatments. These colors were counterbalanced across dyads. This procedure ensured that the doctors would believe that this experiment followed a single-blind procedure, and the patients were the only ones who had no prior knowledge about the color of the Thermedol treatment.

After introducing the two treatments to the doctors, a head-mounted camera recording system was attached to the doctors' head in order to record their facial expressions during the whole experiment. This camera recording system includes a GoPro camera attached to lightweight head-gear [46] and is designed to prevent head and body motion without blocking eye contact in experiments involving social interaction. After setting up the camera, the doctors were told that they would receive a series of pain stimulations to four sites on the volar surface of their left inner forearms. These four sites were organized in a 2 x 2 layout, with one treatment being placed on two sites (e.g. Thermedol on site 1 and 2), and the other treatment being placed on the other two sites (e.g. Control on site 3 and 4). The doctor was instructed that they would receive two pain stimulations on each site, resulting in a total of eight stimulations across all four sites.

Before applying the two treatments, the doctors also received a series of practice trials with one stimulation on each site, totaling four stimulations. Each stimulation was delivered using a 30 by 30 mm ATS thermode (Medoc Ltd, Ramat Yishai, Israel) and lasted about 13 seconds for a higher temperature (47 Celsius), including 1.5 seconds for ramp up, 10 seconds at 47 Celsius, and 1.5 seconds for ramp down. For a lower temperature (43 Celsius), it lasted about 12.4 seconds, including 1.2 seconds for ramp up, 10 seconds at 43 Celsius, and 1.2 seconds for ramp down. During each stimulation, the doctor was first asked to continuously report for 16 seconds how much pain they were experiencing on a 100-point visual analog scale (VAS), where 0 was "no pain experienced" and 100 was "most pain imaginable". Immediately after reporting the continuous rating of pain, the doctor was then asked to report how much pain they experienced overall on a 100-point visual-analog-scale (VAS). Thus, two ratings, namely a continuous and an overall pain rating, were recorded for each stimulation. The codes for running the whole experiment were in Matlab R2015. The experimenter instructed each participant that these practice trials would help familiarize them with the pain reporting procedure. However, they also helped reduce site-specific habituation effects in our data based on results described by Jepma et al., [85], the effectiveness of this procedure can be seen in Extended Data Figure 7(A).

Since the purpose of this phase was to convince the doctor that the Thermedol treatment was more effective in reducing thermal pain than the Control treatment, the experimenter delivered a higher temperature (47 Celsius) to sites with the Control and a lower temperature (43 Celsius) to sites with the Thermedol treatment. Because several placebo studies have demonstrated that healthy participants only exhibit a placebo effect when a control treatment is administered first [43,77,78], the delivery order was always one stimulation on each site with the Control, followed by two stimulations on each site with the Thermedol, and ending with one stimulation on each site with the Control treatment (Figure 1(A-1); e.g. stimulation site: 3-4-2-1-1-2-4-3). Each pain stimulation lasted for 13 seconds, and the doctors were also asked to report the same continuous and overall pain rating used in the practice trials. Additionally, before the very first stimulation, the doctor reported how effective they thought the Thermedol treatment would be in the following stimulations on a 100-point VAS. After the final last stimulation, the doctor reported how effective they found Thermedol to be on a similar scale. Based on these two ratings of effectiveness, we were able to examine whether the manipulation of doctors' beliefs worked in the expected direction; specifically, whether the doctors believed that the Thermedol treatment reduced more pain than the Control treatment.

Doctor-Patient Interaction Phase

The goal of the Doctor-Patient Interaction phase was to examine whether a doctor's beliefs about the treatment impacted a patient's subjective and objective experience of pain. Thus, patients' continuous and overall pain ratings, as well as their facial expressions were recorded. The camera recording systems were attached to patients' heads before they entered the experimental room. Patients were instructed to sit next to the doctors at a 90-degree angle, and a separate computer screen was placed in front of each them. This arrangement ensured that neither the doctor nor the patient could see what the other reported on the screen during the entirety of the interaction

phase. Doctors read a script explaining the experimental procedures to the patients, which included: (1) applying four practice trials, (2) administering the first treatment, (3) completing pre-stimulation ratings, (4) pain stimulation with the first treatment, (5) completing post-stimulation ratings, (6) five minutes of rest, (7) applying four practice trials, (8) administering the second treatment, (9) completing pre-stimulation ratings, (10) pain stimulations with the second treatment, and (11) completing post-stimulation ratings. One of the primary differences between the Doctor Conditioning phase and Doctor-Patient Interaction phase was that the doctor, not the experimenter, led the progress of the experiment.

Similar to practice trials administered during the Doctor Conditioning phase, patients also received one pain stimulation on each of the four sites on their left inner forearms, which served to reduce within site habituation effects (see Extended Data Figure 7(A)). Each stimulation lasted for 13 seconds, including 1.5 seconds for ramp up, 10 seconds at peak temperature (i.e. 47 Celsius), and 1.5 seconds for ramp down. During each stimulation, the patients reported continuous and overall pain ratings on a 100-point VAS. Right after the four practice trials, doctors adhered to the instructions shown on their computer screen and administered the first treatment to the patients. Importantly, doctors did not know which treatment to administer first until this moment. This procedure ensured that the doctors did not reveal their beliefs before administering the first treatment, although based on our design, the first treatment was always the Control treatment. The doctors then used a 1/8 teaspoon to retrieve the Control treatment cream from one of the two containers and applied the cream across the four sites. Since the doctors followed the single-blind procedure, the patients had no prior knowledge which cream was the "real" Thermedol treatment. This part assured that patients' experiences of the two treatments were purely based on their interaction with the doctors.

Before testing the effectiveness of the first treatment, patients were instructed not only to report how effective they thought the first treatment would be in the following stimulations, but also to report their impressions about the doctors, including ratings of empathy, competence, likability, and trust on a 100-point VAS. These pre-stimulation ratings, which served as baseline measurements, were compared with post-stimulation ratings to examine changes in beliefs after the Control treatment. Doctors were also asked to report how effective they thought the Control treatment would be in the following stimulations, which served as a manipulation check that their beliefs did not change from the Conditioning phase to the Interaction phase.

Since the purpose of the Doctor-Patient Interaction phase was to examine whether doctors' beliefs about the treatment impacted patients' subjective and objective experience of pain, the main difference compared to the Doctor Conditioning phase was that the intensity of pain stimulation (i.e. 47 Celsius) was identical to the first and the second treatments. For the first treatment, the patients received a series of pain stimulations with two stimulations on each site, resulting in a total of eight stimulations across the four sites. The delivery order of these stimulations was randomized, and the continuous and overall pain ratings of the patients were recorded during each stimulation. After the last stimulation, the patients also completed their post-stimulation ratings, including beliefs of effectiveness of the first treatment and impressions of the

doctors. The doctors also reported how effective they thought the Control treatment was during the previous stimulations.

After the first treatment, the doctors and patients were asked to fill out a questionnaire (i.e. Interpersonal Reactivity Index [86]), which occupied them during a 5-minute rest period between the two treatments. We did not analyze this questionnaire because not all participants finish all the items because everyone's speed was different. This rest period was intended to prevent a carry-over habituation effect in the second treatment. Immediately after the rest period, the same procedures from the first treatment were repeated for the second treatment, the Thermedol treatment. Pre-stimulation ratings were collected and compared with post-stimulation ratings in order to examine the changes in patients' beliefs of treatment effectiveness as well as their impressions of the doctors. In addition, doctors' beliefs of treatment effectiveness before and after administering Thermedol were also collected. Both the doctors and patients were debriefed at the end of the Doctor-Patient Interaction phase. Before debriefing, to ensure that our results were not influenced by participants' preconceptions about our experiments, we asked participants to write down the purpose of our experiment based on their points of view. We do not believe that our results can be explained by demand characteristics as none of the participants' responses closely approximated our hypotheses. After debriefing, participants were given the opportunity to revisit their informed consent and could freely determine whether to withdraw their data. None of the participants decided to withdraw their data.

Study 2

In Study 2, we first wanted to ensure the robustness of our effects by replicating the experiment in an independent sample (86 participants; 43 dyads). No statistical methods were used to pre-determine sample sizes but our sample sizes were similar to those reported in a previous study [43], which employed a similar experimental design. Second, we wanted to examine order effects and rule out the possibility that our Thermedol effects were not a result of participants habituating to the pain stimulation in a between-subject design. As such, we included two orders of treatment delivery: the original order (i.e. Control first, Thermedol second; Figure 3(B-1)) used in Study 1 and the reverse order (i.e. Thermedol first, Control second; Figure 3(C-1)). Thus, Study 2 also involved the same three phases as Study 1, with the main difference between Study 1 and 2 being that in the Doctor-Patient Interaction phase, we randomized the delivery order of treatments across dyads. Additionally, we examined the effect of both temperature intensity and the experimenter by including two different temperatures (i.e. 46 Celsius and 47 Celsius) and two experimenters in Study 2 (both male).

Study 3

The main purpose of Study 3 (n = 60 participants; 65.00% female) was to address several remaining questions arising from the previous two studies. First, although findings from Study 2 suggested that this effect was unlikely to be explained by habituation in a between-subject design, a within subject design would provide an even stronger test. Thus, we used a within-subject ABBA design (administration order: Control - Thermedol - Thermedol - Control) in the Doctor-Patient

interaction phase. Second, to confirm that our findings were not impacted by the presence of cameras, we removed head-mounted cameras in this study. Third, to fully rule out the alternative hypothesis that the transmission of placebo effect was due to the experimenter, but not the doctor, we included three experimenters (experimenter 1, 2, and 3) in Study 3. During the doctor conditioning phase, experimenter 1 trained the doctor and created the creams, but experimenter 1 left the experimental room at the end of this phase. Experimenter 2 and 3 alternatively run the doctor-patient interaction phase across dyads, and they were completely blind to which cream contained the Thermedol. This procedure ensured that the doctor was the only person in the experimental room who knew which cream contained the active treatment. In addition, in order to examine whether this effect could be replicated in a higher temperature, we used a higher temperature in this study (48 Celsius).

Skin Conductance Response (SCR) recording and preprocessing

We measured SCRs at the index and middle fingers of the left hand by using EDA electrodes (EL507, Biopac Systems, Goleta, CA, USA) and 0.5%-NaCl electrode paste (GEL101, Biopac Systems). The SCRs were recorded by using a BIOPAC MP150 system and Acknowledge software. In Study 1, the sampling rate was 200Hz, and data were first filtered with a 1st order Butterworth filter with a 5Hz low pass filter and 0.01 Hz high pass filter. The filtered SCR data was then down-sampled to 1 Hz. In Study 2, besides of the sampling rate, which was 2000 Hz, all the other preprocessing steps were identical to Study 1. In Study 3, since the duration of the whole experiment was longer than the previous studies, we followed the preprocessing steps as identical to those used in Study 1 for each block individually and normalized the SCRs within each block. All of the preprocessing steps were done within Python 2.7 and Python 3.6. The Python packages used in preprocessing and analysis including Pandas [87], Numpy [88], Seaborn [89], Matplotlib [90], Scikit-learn [91], Scipy [92,93], and Nltools [94].

In order to match the length of continuous pain ratings, we extracted patients' SCRs starting from the onset of each pain stimulation for 16 seconds. Two out of 24 patients from Study 1,13 out of 43 patients from Study 2, and 6 out of 30 patients in Study 3 were excluded due to minimal SCRs during pain stimulations (i.e., non-responders). In addition, we were unable to analyze missing data from one patient from Study 1 and four trials in two patients from Study 3. We then computed the area-under-curve (AUC) for each pain stimulation using the AUC function from Scikit-learn [91]. The AUC represented the overall SCR amplitude in each pain stimulation and was then used in further mixed-effect generalized linear model (GLM) analyses.

Analyses

Linear mixed-effects models with the R lme4 package [95], lmerTest [96], lsmeans [97], and the Pymer4 python to R interface [98] were used to examine whether when the Thermedol and Control treatments were administered, participants reported or experienced differences in their beliefs of effectiveness as well as pain ratings and SCRs. Across all analyses, subjects were treated as having random intercepts and slopes with respect to trial numbers (1 to 8), time (pre- or post-

stimulation) as well as condition (Thermedol or Control). In Study 1, to examine whether during the Doctor Conditioning phase, doctors reported differences in beliefs of effectiveness between the two treatments, we used a linear mixed-effects model with the condition (Thermedol – Control) as a fixed effects factor (Extended Data Figure 1). To examine whether during the Doctor-Patient Interaction phase, doctors or patients reported changes in beliefs of effectiveness, we used a linear mixed-effects model with condition (Thermedol – Control) by time (Post – Pre) as an interaction term (Supplementary Table 1). The same model was also used to test whether patients reported finding doctors differing in empathy between the two treatments (Supplementary Table 1).

To test whether doctors or patients experienced differences in thermal pain (overall pain rating) between the two treatments, we used a linear mixed-effects model with condition (Thermedol – Control) as a factor, and with stimulation sites (1 to 4), colors of treatment creams (red or blue), and trial numbers included as covariates of no interests (Supplementary Table 1). We were unable to analyze missing pain rating data from one doctor in Study 1. We also used the same model to explore whether patients experienced differences in thermal pain (maximal pain ratings) between the two treatments (Supplementary Table 1). To examine whether patients also showed differences in SCR to pain stimulations (AUC), one of the objective measures of pain experiences, we also used a linear mixed-effects model with condition (Thermedol – Control) as a factor, and with the same covariates (Supplementary Table 1). We lastly used the same model to test differences in doctors' or patients' pain facial expressions (predicted pain ratings) between the two treatments (Supplementary Table 1).

In Study 2, since we counterbalanced the order (original or reverse) across participants, we used a linear mixed-effects model with condition (Thermedol – Control) by time (Post – Pre) by order (original – reverse) as an interaction term to explore whether patients experienced differences in thermal pain (overall pain rating) between the two treatments. In this model, stimulation sites (1 to 4), colors of treatment creams (red or blue), trial numbers, temperature (46 or 47 Celsius), and experimenter (1 or 2) were included as covariates of no interests (Supplementary Table 3). The same model was also used to analyze whether patients showed differences in SCR to pain stimulations (AUC) between the two treatments (Supplementary Table 3). To examine whether during the Doctor-Patient Interaction phase, doctors or patients reported changes in beliefs of effectiveness, we used a linear mixed-effects model with condition (Thermedol – Control) by time (Post – Pre) by order (original – reverse) as an interaction term (Supplementary Table 3).

In Study 3, to test whether doctors or patients experienced differences in thermal pain (overall pain rating) between the two treatments, we used a linear mixed-effects model with condition (Thermedol – Control) as a factor, and with stimulation sites (1 to 4), colors of treatment creams (red or blue), and trial numbers included as covariates of no interest (Supplementary Table 4). To examine whether patients also showed differences in SCR to pain stimulations (AUC), one of the objective measures of pain experiences, we also used a linear mixed-effects model with condition (Thermedol – Control) as a factor, and with the same covariates plus a linear effect of time (Supplementary Table 4).

Facial Expressions

Facial behaviors during the experiment were recorded using GoPro HERO 4 cameras recording at 120 frames per second at 1920 x 1080 resolution. Each camera was positioned using a head-mounted fixture [46] that would provide a consistent view of the face without obstructing the participant's view. Recorded videos were then aligned to stimuli onsets using audio triggers. Facial behavior features consisting of 20 facial action units (AU), a standard for measuring facial muscle movement based on the Facial Action Coding System [FACS, 49], plus three rotational orientations of head (pitch, yaw, and roll) were extracted using the FACET algorithm [47] accessed through the iMotions biometric research platform [48]. The median value was subtracted from each feature to control for baseline variability and all features were downsampled to 10 hz. For each trial, sixteen seconds following stimulation onset which included the ramp up and ramp down times for stimulation was considered for analysis. Due to the technical difficulties from Study 2 in retrieving accurate onset time of each pain stimulation, we were only able to analyze facial expression data from Study 1 (Face expressions were intentionally not recorded in Study 3).

Pain facial expression model
From the timeseries of facial movement for each trial we extracted three facial activity descriptors consisting of the probability of maximal expression (max), probability of minimal expression (min), and time to maximal expression (tmax) of the original time-series data (Werner et al, 2017 IEEE). Using these features, we predicted subject-wise standardized pain ratings across the 12 trials in the doctor's conditioning phase using a linear ridge regression with a nested 5 fold cross validation for hyper-parameter optimization. We evaluated the model using a leave-one-subject-out (LOSO) cross validation which yielded an average within-subject Pearson correlation between predicted and actual pain ratings of $r = .41$ ($sd = .33$), $p = .003$ (Extended Data Figure 3(C)). The full model trained on the entire conditioning phase data was then used to predict pain ratings from the Patient's face during the interaction phase which yielded an average within-subject Pearson correlation of $r = .24$ ($sd = .31$), $p = .003$. Inference on model performance was determined via Monte-Carlo based permutation testing, in which the training labels were randomly shuffled and the model was retrained 5,000 times (Extended Data Figure 3(B)). Top features included activities related to AU4 (brow lowerer), pitch, and AU12 (lip corner puller) (Extended Data Figure 3(A)). Predicted pain ratings from the patient's facial expressions were compared between the Placebo and the Thermedol conditions using a linear mixed-effects model including both condition and trial numbers as random effects.

Face Model Visualization
We used our Python Facial Expression Analysis Toolbox (version 0.0.1) (Cheong, Byrnes, and Chang, unpublished manuscript) to visualize how specific features from our model correspond to changes in facial morphometry. In brief, we learned a mapping between 20 facial action unit intensities and 68 landmarks comprising a 2-dimensional face using partial least squares implemented in scikit-learn [91]. We used 10,708 images corresponding to each frame from the extended cohn-kanade facial expression database [CK+, 99] and first extracted the landmarks using OpenFace [100] and action units from iMotions FACET engine [48]. Next, we performed an affine transformation to a canonical face prior to fitting the model and added pitch, roll, and yaw head

rotation parameters as covariates with the 20 action unit features. We then fit our model using a 3-Fold cross-validation scheme and achieved an overall training r^2 = .61 and a cross-validated mean r^2 = .53. We used this facial morphometry model to visualize the positive beta weights corresponding to the maximum intensity descriptors from the pain expression model scaled by 200. This allows us to provide several different visualizations of our pain model. First, we can plot the vector fields of how each facial landmark changes from a neutral facial expression to our pain facial expression model. Second, we can also plot the parameter estimates of the model corresponding to each action unit. Stronger intensities indicate a greater contribution to the overall prediction of subject pain. We note that these coefficients reflect a subset of the complete model which includes features difficult to visualize such as temporal information (time to maximum). Results are shown in Figure 2(B-1) and Extended Data Figure 3(A).

Data availability

The data that support the findings of this study are available in Github: https://github.com/cosanlab/socially_transmitted_placebo_effects/

Code availability

The code that support the findings of this study are available in Github: https://github.com/cosanlab/socially_transmitted_placebo_effects/

References

1. Beecher, H. K. The powerful placebo. *J. Am. Med. Assoc.* **159**, 1602–1606 (1955).
2. Gold, H., Kwit, N. T. & Otto, H. The xanthines (theobromine and aminophylline) in the treatment of cardiac pain. *JAMA* **108**, 2173–2179 (1937).
3. Rosenzweig, S. Some implicit common factors in diverse methods of psychotherapy. *Am. J. Orthopsychiatry* **6**, 412–415 (1936).
4. Houston, W. R. The doctor himself as a therapeutic agent. *Ann. Intern. Med.* **11**, 1416 (1938).
5. Uhlenhuth, E. H., Canter, A., Neustadt, J. O. & Payson, H. E. The symptomatic relief of anxiety with meprobamate, phenobarbital and placebo. *Am. J. Psychiatry* **115**, 905–910 (1959).
6. Shapiro, A. K. & Shapiro, E. *The powerful placebo: From ancient priest to modern physician*. (JHU Press, 2000).
7. Ioannidis, J. P. *et al.* Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* **286**, 821–830 (2001).
8. Świder, K. & Bąbel, P. The effect of the sex of a model on nocebo hyperalgesia induced by social observational learning. *PAIN®* **154**, 1312–1317 (2013).
9. Vögtle, E., Barke, A. & Kröner-Herwig, B. Nocebo hyperalgesia induced by social observational learning. *Pain* **154**, 1427–1433 (2013).
10. Colloca, L. & Benedetti, F. Placebo analgesia induced by social observational learning. *Pain* **144**, 28–34 (2009).
11. Koban, L. & Wager, T. D. Beyond conformity: Social influences on pain reports and physiology. *Emotion* **16**, 24–32 (2016).
12. Yoshida, W., Seymour, B., Koltzenburg, M. & Dolan, R. J. Uncertainty increases pain: evidence for a novel mechanism of pain modulation involving the periaqueductal gray. *J. Neurosci.* **33**, 5638–5646 (2013).
13. Haaker, J., Yi, J., Petrovic, P. & Olsson, A. Endogenous opioids regulate social threat learning in humans. *Nat. Commun.* **8**, 15495 (2017).
14. Benedetti, F., Durando, J. & Vighetti, S. Nocebo and placebo modulation of hypobaric hypoxia headache involves the cyclooxygenase-prostaglandins pathway. *Pain* **155**, 921–928 (2014).
15. Holtzheimer, P. E. *et al.* Subcallosal cingulate deep brain stimulation for treatment-resistant depression: a multisite, randomised, sham-controlled trial. *The lancet. Psychiatry* **4**, 839–849 (2017).
16. Ashar, Y. K., Chang, L. J. & Wager, T. D. Brain Mechanisms of the Placebo Effect: An Affective Appraisal Account. *Annu. Rev. Clin. Psychol.* **13**, 73–98 (2017).
17. Price, D. D., Finniss, D. G. & Benedetti, F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu. Rev. Psychol.* **59**, 565–590 (2008).
18. Benedetti, F., Mayberg, H. S., Wager, T. D., Stohler, C. S. & Zubieta, J.-K. Neurobiological mechanisms of the placebo effect. *J. Neurosci.* **25**, 10390–10402 (2005).
19. Benedetti, F. *The Patient's Brain: The Neuroscience Behind the Doctor-patient Relationship*. (OUP Oxford, 2011).
20. Geuter, S., Koban, L. & Wager, T. D. The Cognitive Neuroscience of Placebo Effects: Concepts, Predictions, and Physiology. *Annu. Rev. Neurosci.* **40**, 167–188 (2017).
21. de la Fuente-Fernández, R. *et al.* Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science* **293**, 1164–1166 (2001).
22. Jensen, K. B. *et al.* Sharing pain and relief: neural correlates of physicians during treatment of patients. *Mol. Psychiatry* **19**, 392–398 (2014).
23. Olanow, C. W. *et al.* Gene delivery of neurturin to putamen and substantia nigra in

Parkinson disease: A double-blind, randomized, controlled trial. *Ann. Neurol.* **78**, 248–257 (2015).

24. Luborsky, L. *et al.* The researcher's own therapy allegiances: A 'wild card' in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice* **6**, 95–106 (1999).
25. Walach, H., Sadaghiani, C., Dehm, C. & Bierman, D. The therapeutic effect of clinical trials: understanding placebo response rates in clinical trials--a secondary analysis. *BMC Med. Res. Methodol.* **5**, 26 (2005).
26. Greenberg, R. P., Bornstein, R. F., Zborowski, M. J., Fisher, S. & Greenberg, M. D. A meta-analysis of fluoxetine outcome in the treatment of depression. *J. Nerv. Ment. Dis.* **182**, 547–551 (1994).
27. Holroyd, K. A., Tkachuk, G., O'Donnell, F. & Cordingley, G. E. Blindness and bias in a trial of antidepressant medication for chronic tension-type headache. *Cephalalgia* **26**, 973–982 (2006).
28. Margraf, J. *et al.* How 'Blind' Are Double-Blind Studies? *J. Consult. Clin. Psychol.* **59**, 184–187 (1991).
29. Morin, C. M. *et al.* How Blind Are Double-Blind Placebo-Controlled Trials of Benzodiazepine Hypnotics. *Sleep* **18**, 240–245 (1995).
30. Miller, F. G., Colloca, L. & Kaptchuk, T. J. The placebo effect: illness and interpersonal healing. *Perspect. Biol. Med.* **52**, 518–539 (2009).
31. Blasini, M., Peiris, N., Wright, T. & Colloca, L. The Role of Patient-Practitioner Relationships in Placebo and Nocebo Phenomena. *Int. Rev. Neurobiol.* **139**, 211–231 (2018).
32. Rosenthal, R. & Rubin, D. B. Interpersonal expectancy effects: the first 345 studies. *Behav. Brain Sci.* **1**, 377–386 (1978).
33. Rosenthal, R. Interpersonal Expectancy Effects: A 30-Year Perspective. *Curr. Dir. Psychol. Sci.* **3**, 176–179 (1994).
34. Pfungst, O. *Clever Hans:(the horse of Mr. Von Osten.) a contribution to experimental animal and human psychology*. (Holt, Rinehart and Winston, 1911).
35. Rosenthal, R. & Lawson, R. A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats. *J. Psychiatr. Res.* **2**, 61–72 (1964).
36. Rosenthal, R. & Jacobson, L. Pygmalion in the classroom. *Urban Rev.* **3**, 16–20 (1968).
37. Rosenthal, R. *Experimenter effects in behavioral research*. (Appleton-Century-Crofts, 1966).
38. Doyen, S., Klein, O., Pichon, C.-L. & Cleeremans, A. Behavioral priming: it's all in the mind, but whose mind? *PLoS One* **7**, e29081 (2012).
39. Joyce, A. S. & Piper, W. E. Expectancy, the therapeutic alliance, and treatment outcome in short-term individual psychotherapy. *J. Psychother. Pract. Res.* **7**, 236–248 (1998).
40. Meyer, B. *et al.* Treatment expectancies, patient alliance and outcome: Further analyses from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J. Consult. Clin. Psychol.* **70**, 1051–1055 (2002).
41. Arnkoff, D. B., Glass, C. R. & Shapiro, S. J. Expectations and preferences. in *Psychotherapy relationships that work; therapist contributions and responsiveness to patients* 325–346 (psycnet.apa.org, 2002).
42. Gracely, R. H., Dubner, R., Deeter, W. R. & Wolskee, P. J. Clinicians Expectations Influence Placebo Analgesia. *Lancet* **1**, 43–43 (1985).
43. Schafer, S. M., Colloca, L. & Wager, T. D. Conditioned placebo analgesia persists when subjects know they are receiving a placebo. *J. Pain* **16**, 412–420 (2015).
44. Voudouris, N. J., Peck, C. L. & Coleman, G. The role of conditioning and verbal expectancy in the placebo response. *Pain* **43**, 121–128 (1990).
45. Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *Am. Psychol.* **17**, 776 (1962).

46. Cheong, J., Brooks, S. & Chang, L. J. FaceSync: Open source framework for recording facial expressions with head-mounted cameras. *F1000Res.* **8**, (2019).

47. Littlewort, G. *et al.* The computer expression recognition toolbox (CERT). in *Face and Gesture 2011* 298–305 (2011).

48. *iMotions Biometric Research Platform 6.0*. (iMotions A/S, Copenhagen, Denmark, 2016).

49. Ekman, P. & Friesen, W. V. Measuring facial movement. *J. Nonverbal Behav.* **1**, 56–75 (1976).

50. Prkachin, K. M. The consistency of facial expressions of pain: a comparison across modalities. *Pain* **51**, 297–306 (1992).

51. Lucey, P. *et al.* Automatically Detecting Pain in Video Through Facial Action Units. *IEEE Trans. Syst. Man Cybern. B Cybern.* **41**, 664–674 (2011).

52. Walach, H., Schimdt, S., Bihr, Y.-M. & Wiesch, S. The effects of a caffeine placebo and experimenter expectation on blood pressure, heart rate, well-being, and cognitive performance. *Eur. Psychol.* **6**, 15 (2001).

53. Walach, H., Schmidt, S., Dirhold, T. & Nosch, S. The effects of a caffeine placebo and suggestion on blood pressure, heart rate, well-being and cognitive performance. *Int. J. Psychophysiol.* **43**, 247–260 (2002).

54. Halverson, A. M., Hallahan, M., Hart, A. J. & Rosenthal, R. Reducing the biasing effects of judges' nonverbal behavior with simplified jury instruction. *J. Appl. Psychol.* **82**, 590 (1997).

55. Learman, L. A., Avorn, J., Everitt, D. E. & Rosenthal, R. Pygmalion in the nursing home. The effects of caregiver expectations on patient outcomes. *J. Am. Geriatr. Soc.* **38**, 797–803 (1990).

56. Wu, L. M., Mohamed, N. E., Winkel, G. & Diefenbach, M. A. Patient and spouse illness beliefs and quality of life in prostate cancer patients. *Psychol. Health* **28**, 355–368 (2013).

57. Wampold, B. E., Mondin, G. W., Moody, M. & Stich, F. A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially,' all must have prizes.' *Psychological* (1997).

58. Wampold, B. E. & Imel, Z. E. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. (Routledge, 2015).

59. Weinberger, J. Common Factors Aren't So Common: The Common Factors Dilemma. *Clinical Psychology: Science and Practice* **2**, 45–69 (1995).

60. Frank, J. D. & Frank, J. B. *Persuasion and healing: A comparative study of psychotherapy*. (JHU Press, 1993).

61. Rogers, C. R. The necessary and sufficient conditions of therapeutic personality change. *J. Consult. Psychol.* **21**, 95–103 (1957).

62. Yahne, C. E. & Miller, W. R. Evoking hope. in *Integrating spirituality into treatment: Resources for practitioners.* (ed. Miller, W. R.) 217–233 (American Psychological Association, 1999).

63. Ackerman, S. J. & Hilsenroth, M. J. A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clin. Psychol. Rev.* **23**, 1–33 (2003).

64. Lester, G. W. & Smith, S. G. Listening and talking to patients. A remedy for malpractice suits? *West. J. Med.* **158**, 268–272 (1993).

65. Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P. & Auerbach, A. Therapist success and its determinants. *Arch. Gen. Psychiatry* **42**, 602–611 (1985).

66. Okiishi, J., Lambert, M. J., Nielsen, S. L. & Ogles, B. M. Waiting for supershrink: an empirical analysis of therapist effects. *Clin. Psychol. Psychother.* **10**, 361–373 (2003).

67. Wampold, B. E. & Brown, G. S. J. Estimating variability in outcomes attributable to therapists: a naturalistic study of outcomes in managed care. *J. Consult. Clin. Psychol.* **73**, 914–923 (2005).

68. Martin, D. J., Garske, J. P. & Davis, M. K. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *J. Consult. Clin. Psychol.* **68**, 438–450 (2000).

69. Derksen, F., Bensing, J. & Lagro-Janssen, A. Effectiveness of empathy in general practice: a systematic review. *Br. J. Gen. Pract.* **63**, e76–84 (2013).

70. Horvath, A. & Luborsky, L. The role of the therapeutic alliance in psychotherapy. *J. Consult. Clin. Psychol.* **61**, 561–573 (1993).

71. Horvath, A. & Symonds, B. D. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *J. Couns. Psychol.* **38**, 139 (1991).

72. Rakel, D. *et al.* Perception of empathy in the therapeutic encounter: effects on the common cold. *Patient Educ. Couns.* **85**, 390–397 (2011).

73. Necka, E. A. & Atlas, L. Y. The Role of Social and Interpersonal Factors in Placebo Analgesia. in *International Review of Neurobiology* (Elsevier, 2018).

74. Thurstone, L. L. A law of comparative judgment. *Psychol. Rev.* **34**, 273 (1927).

75. Fechner, G. T. Elemente der psychophysik (2 vols). breitkopf and hartel. vol. 1 trans, by he adler (1966). *Elements of psychophysics* (1860).

76. Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47**, 263–291 (1979).

77. Flaten, M. A., Aslaksen, P. M., Lyby, P. S. & Bjørkedal, E. The relation of emotions to placebo responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 1818–1827 (2011).

78. Lyby, P. S., Aslaksen, P. M. & Flaten, M. A. Is fear of pain related to placebo analgesia? *J. Psychosom. Res.* **68**, 369–377 (2010).

79. Wager, T. D. & Atlas, L. Y. The neuroscience of placebo effects: connecting context, learning and health. *Nat. Rev. Neurosci.* **16**, 403–418 (2015).

80. Batt-Rawden, S. A., Chisolm, M. S., Anton, B. & Flickinger, T. E. Teaching empathy to medical students: an updated, systematic review. *Acad. Med.* **88**, 1171–1177 (2013).

81. Sulzer, S. H., Feinstein, N. W. & Wendland, C. L. Assessing empathy development in medical education: a systematic review. *Med. Educ.* **50**, 300–310 (2016).

82. Fine, V. K. & Therrien, M. E. Empathy in the doctor-patient relationship: skill training for medical students. *J. Med. Educ.* **52**, 752–757 (1977).

83. Ha, J. F. & Longnecker, N. Doctor-patient communication: a review. *Ochsner J.* **10**, 38–43 (2010).

84. Vandewauw, I. *et al.* A TRP channel trio mediates acute noxious heat sensing. *Nature* **555**, 662–666 (2018).

85. Jepma, M., Jones, M. & Wager, T. D. The dynamics of pain: evidence for simultaneous site-specific habituation and site-nonspecific sensitization in thermal pain. *J. Pain* **15**, 734–746 (2014).

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

87. McKinney, W. & Others. Data structures for statistical computing in python. in *Proceedings of the 9th Python in Science Conference* **445**, 51–56 (Austin, TX, 2010).

88. Oliphant, T. E. *A guide to NumPy*. **1**, (Trelgol Publishing USA, 2006).

89. Waskom, M. *et al.* Seaborn: statistical data visualization. *URL: https://seaborn. pydata. org/(visited on 2017-05-15)* (2014).

90. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

91. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

92. Oliphant, T. E. Python for Scientific Computing. *Computing in Science Engineering* **9**, 10–20 (2007).

93. Millman, K. J. & Aivazis, M. Python for Scientists and Engineers. *Computing in Science Engineering* **13**, 9–12 (2011).

94. Chang, L., Jolly, E., Cheong, J. H., Burnashev, A. & Chen, P.-H. A. *cosanlab/nltools: 0.3.11*. (2018). doi:10.5281/zenodo.2229813

95. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using

lme4. *arXiv [stat.CO]* (2014).

96. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest: Tests for random and fixed effects for linear mixed effect models. R package version 2.0-11. *URL http://CRAN. R-project. org/package= lmerTest* (2014).

97. Lenth, R. V. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw.* **69**, 1–33 (2016).

98. Jolly, E. Pymer4: Connecting R and Python for Linear Mixed Modeling. *The Journal of Open Source Software* **3**, 862 (2018).

99. Lucey, P. *et al.* The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops* 94–101 (ieeexplore.ieee.org, 2010).

100. Baltrušaitis, T., Robinson, P. & Morency, L. P. OpenFace: An open source facial behavior analysis toolkit. in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* 1–10 (ieeexplore.ieee.org, 2016).

Acknowledgments

Author contributions

All authors designed the study. P.-H.C., J.H.C., E.J. and H.E., collected the data. P.-H.C., J.H.C. and L.J.C. analyzed the data. P.-H.C., J.H.C., E.J., T.D.W. and L.J.C. wrote the paper.

Competing interests

The authors declare no competing interests.