



OPEN

A virtual experimenter does not increase placebo hypoalgesia when delivering an interactive expectancy manipulation

Bjoern Horing^{1✉}, Sarah C. Beadle², Zachariah Inks³, Andrew Robb³, Eric R. Muth^{2,4} & Sabarish V. Babu³

Lack of standardization and unblinding threaten the research of mechanisms involved in expectancy effects on pain. We evaluated a computer-controlled virtual experimenter (VEx) to avoid these issues. Fifty-four subjects underwent a baseline-retest heat pain protocol. Between sessions, they received an expectancy manipulation (placebo or no-treatment) delivered by VEx or text-only control condition. The VEx provided standardized “social” interaction with the subjects. Pain ratings and psychological state/trait measures were recorded. We found an interaction of expectancy and delivery on pain improvement following the intervention. In the text conditions, placebo was followed by lower pain, whereas in the VEx conditions, placebo and no-treatment were followed by a comparable pain decrease. Secondary analyses indicated that this interaction was mirrored by decreases of negative mood and anxiety. Furthermore, changes in continuous pain were moderated by expectation of pain relief. However, retrospective pain ratings show an effect of expectancy but not of delivery. We conclude that we successfully applied an automated protocol for inducing expectancy effects on pain. The effect of the VEx regardless of treatment may be due to interactions of attention allocation and locus of control. This points to the diversity of expectancy mechanisms, and has implications for research and computer-based treatment applications.

Pain is a challenge to human well-being owing to its high prevalence as a clinical symptom¹, corresponding societal costs and substantial impact on quality of life². In many painful clinical conditions and experimental pain modalities, placebo effects account for a large amount of symptom change following interventions^{3–5}. Because new treatments perform against placebo conditions, it is desirable that the variance attributable to placebo effects is as low as possible to better detect the efficacy of treatments⁶. Here, we report results using a virtual experimenter (VEx) to standardize interventions to more reliably control variance attributable to placebo effects.

Placebo effects are mediated by expectancies⁵—since symptom changes can be positive or negative (as in nocebo effects), these can more generally be termed “expectancy effects”. Unless a treatment is not perceived⁷, expectancy effects are implicated in *every* treatment, regardless of whether it involves an “active” component^{5,6}. Expectancies arise from the complex interaction of information delivered alongside and through a treatment (for example, assertions of efficacy or potential side effects), explicit preconceptions about that treatment (for example, attitudes towards the treatment modality, such as acupuncture or medication), implicit preconceptions (for example, prior experiences, including conditioning)⁸, as well as mindset, mood and other determinants⁵.

When social interactions are involved, several related concepts, such as sympathy or trust, codetermine treatment outcome to a large degree⁹. For example, a pleasant interaction could per se elevate mood and well-being of a person (direct psychosocial effect), or it could facilitate adherence or the acceptance of treatment instructions (indirect psychosocial effect). Negative expectancies, such as those mediated by doubts about the intentions or the competence of a physician, could reduce or even abolish treatment effects¹⁰, or lead to an increased reporting of side effects¹¹.

¹Affective Neuroscience Group, Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany. ²Department of Psychology, Clemson University, Clemson, SC, USA. ³Division of Human Centered Computing, School of Computing, Clemson University, Clemson, SC, USA. ⁴Division of Research and Economic Development, North Carolina Agricultural and Technical State University, Greensboro, NC, USA. ✉email: b.horing@uke.de

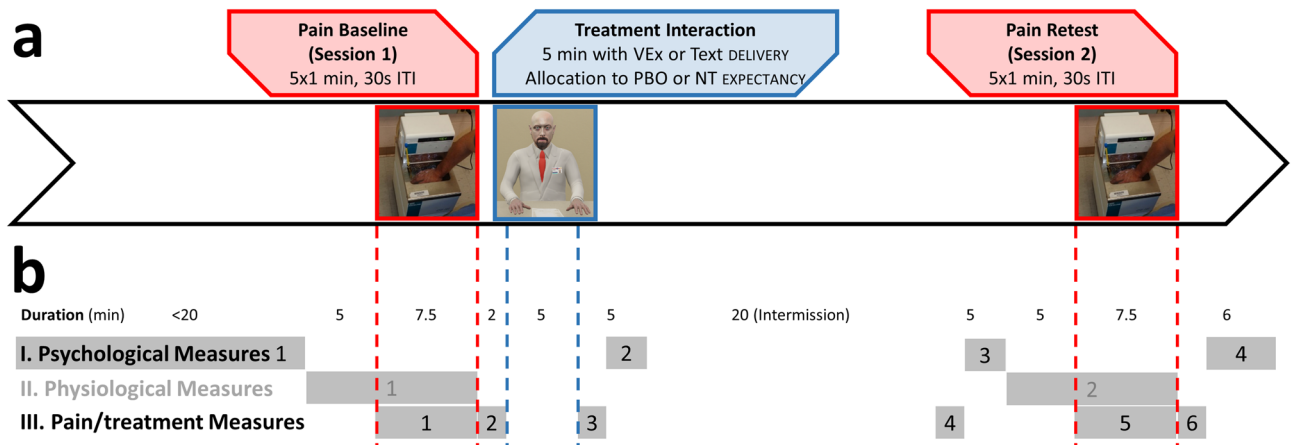


Figure 1. Experimental design and protocol. **(a)** Between pain baseline and retest including 5 hot water immersions, subjects were randomly assigned to DELIVERY and EXPECTANCY conditions. Virtual experimenter screen captured in Unity. **(b)** Timing of measurements. Approximate durations are displayed in the upper row, grey areas indicate the duration of individual measurements, as described below. I. Four psychological measurements, 1 (General Information, IPC, LOT-R, PVAQ, FPQ-III, BFI-44, STAI-Trait, PHQ9, PHQ15, STAI-State 1, MDMQ-SF1), 2 (SPS, NMQ, EIC), 3 (STAI-State 2, MDMQ-SF2), 4 (Manipulation check, exploratory items). II. Physiological measures (data not shown). III. Pain/treatment measures, 1 (continuous ratings), 2 (retrospective ratings), 3 (side effects questionnaire, expectation of pain relief), 4 (side effects questionnaire), 5 (continuous ratings), 6 (retrospective ratings).

Biomedical research, and particularly research on expectancy effects, faces several methodological challenges. Two of these challenges are standardization and replicability^{12,13}, which is difficult to achieve in psychosocial interactions, and blinding¹⁴, that is, that subjects and experimenters are unaware of whether the subject has been assigned to treatment or control group. Failure to address these challenges introduces systematic error that can render experimental results uninterpretable, such as self-fulfilling prophecies and implicit biases^{15–17}. As we have argued before¹⁸, automating the tasks of experimental interactions and treatment allocation with a computerized experimenter could overcome some of these challenges.

Virtual reality simulations have been used as interventions before (e.g. as exposure therapy, or for digital pain management^{19–23}), but not to address the breakpoints involved in human–human interaction especially in clinical studies. Our research indicates that standardization benefits from the introduction of virtual humans, for example in the identification of suspects in police line-ups presented to eye-witnesses, or in training novice healthcare providers in patient interviewing via standardized virtual patients^{24,25}. Here, we followed our research agenda laid out before¹⁸ to explore a virtual experimenter’s (VEx) efficacy to avoid biases during the induction of positive expectancies in subjects undergoing a heat pain protocol, compared with a text/audio control condition. We employed a placebo hypoalgesia paradigm¹⁸, including an inert pill presented as an effective analgesic. We hypothesized that the expectancies should lead to lower pain reports, compared to the no-treatment control. Furthermore, we expected an interaction effect such that in the VEx condition, the impact of psychosocial factors would be more pronounced, facilitating expectancy effects on pain reporting.

Methods

Subjects. Fifty-four healthy volunteers (sex 29f:25m, age 19.6 ± 2.6) were recruited from Clemson University’s student population through word of mouth and an online recruitment system; sample size calculation is provided in Supplementary Text S1. All methods (recruitment and experiment) were carried out in accordance with relevant guidelines and regulations. Subjects filled out an online screening and survey. Inclusion criteria assessed online were: between 18 and 35 years of age, absence of past or present brain or nerve conditions or pain disorders, absence of open wounds or skin conditions on the dominant hand, and abstinence from pain medication and recreational drugs for at least 24 h before the lab visit. Any regular medication was also documented. Furthermore, subjects were only included below the clinical cutoffs provided by the Patient Health Questionnaire for depression ($PHQ9 < 11$) and somatization disorder ($PHQ15 < 11$)^{26,27}.

Following screening, subjects were scheduled for the lab visit. The lab visit included another short survey, two pain measurement sessions (baseline and retest) and the experimental manipulation. Experiments were conducted from August 2017 through February 2018.

Experimental design and overview. The experiment followed a 2×2 between-group design with nested 2×5 within-person repeated pain measurements (Fig. 1a). The repeated measurements were constituted by 5 baseline and 5 retest measurements (repeated measurement factors IMMERSION and SESSION; see “Experimental procedure” and “Pain protocol”). Between the two sessions, subjects were allocated to one of four groups defined by combination of the two between-group factors instruction delivery (factor DELIVERY) and expectancy manipulation (factor EXPECTANCY). DELIVERY refers to the mode of communication employed during instruc-

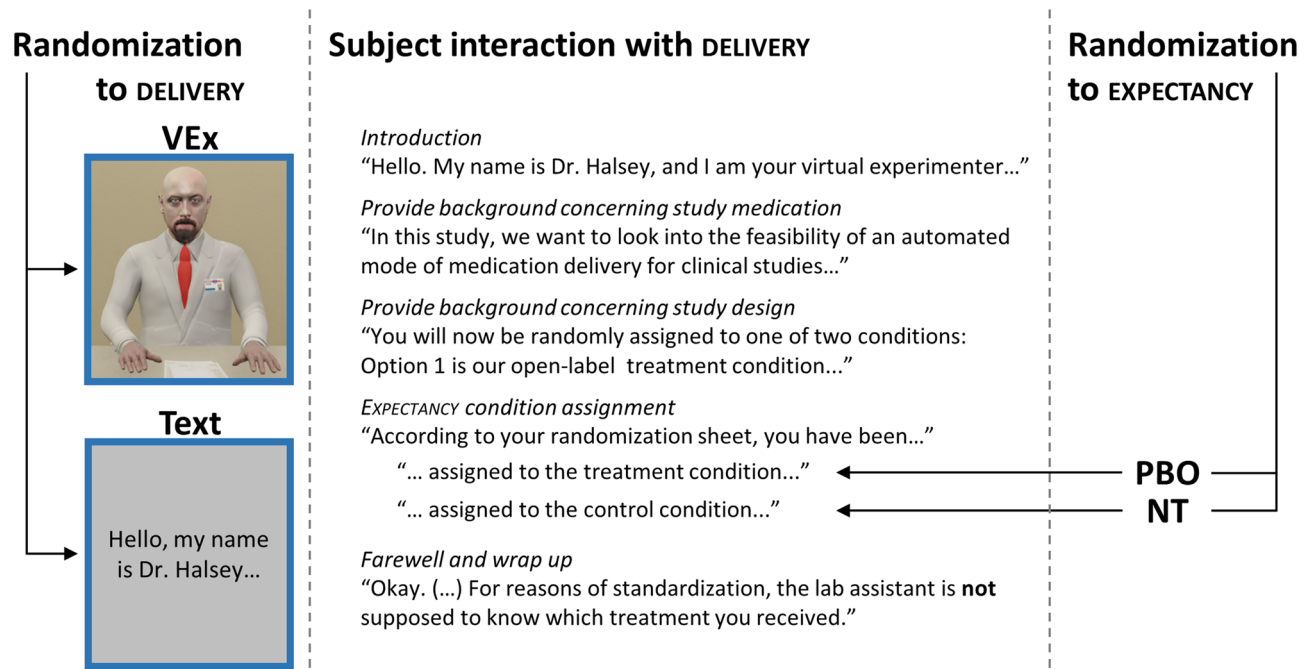


Figure 2. Treatment interaction with virtual experimenter or text/audio condition. Subjects randomly encountered either DELIVERY condition. They then interacted with it (including confirmation of several sections, and option to ask FAQs). During the interaction, they were randomly assigned to either EXPECTANCY condition. Virtual experimenter screen captured in Unity.

tion delivery, either via Text or VEx (both conditions were accompanied by the same audio). EXPECTANCY refers to whether a subject received a treatment introduced as an effective analgesic (a capsule really containing an inert substance, that is, a placebo/PBO), or no treatment (NT).

Experimental procedure. The protocol was reviewed by Clemson University’s Institutional Review Board (vote IRB2016-349), and considered to conform to ethical standards. Subjects signed informed consent prior to participation, which excluded the fact that they might receive a deceptive instruction about their treatment (in PBO EXPECTANCY condition). At conclusion of the study, all subjects were fully debriefed and had the opportunity to withdraw without repercussions. None of the subjects reported being upset after the deception was revealed, and none opted for withdrawal.

Sequence of measurements and timings of the protocol are displayed in Fig. 1b. The experiment lasted about 1.5 h. It was conducted by a female experimenter (SB), in a room maintained at approximately 21 °C. Following instructions, preparations and baseline psychophysiological measurement, the baseline pain measurement started (SESSION 1). Throughout the five one minute hot water immersions, subjects continually rated their pain on a touch screen. After SESSION 1, subjects entered a neighboring room where they received the EXPECTANCY manipulation by allocation to one of four experimental groups on a large screen monitor (see Fig. 2 as well as “Experimental factor 1: instruction delivery” and “Experimental factor 2: expectancy manipulation”).

Importantly, in order to avoid corresponding biases, the human experimenter was unaware of group allocation in both experimental factors, and the subject was asked not to reveal allocation to the human experimenter. All subjects complied with this request, as per experimenter debriefing.

Subjects then returned to the main room for a 30 min intermission. Cover story for the intermission was for the supposed pain medication to take effect; the major intended side benefit was to allow for recovery of skin sensitivity. Subjects were instructed to avoid physical or stressful activities, but could otherwise spend the time ad libitum (e.g., by reading provided magazines). After the intermission, a retest pain measurement was conducted (SESSION 2). Subjects were then debriefed and compensated for participation (\$20, or \$10 and course credit).

Various ratings and questionnaires were administered intermittently (Fig. 1b); heart rate was measured during the entire experiment. Note that psychophysiological measures were considered beyond the scope of this article to allow greater focus on behavioral measures.

Pain protocol. We used a hot water bath (RTE-111, Neslab Instruments, Inc., Newington, NH) as an established method for painful stimulation^{28,29}. The device is capable of circulating the water to avoid areas of cooler water insulating the skin. Subjects were asked to immerse their dominant hand in the water kept at a constant temperature of 47 °C, which is quite painful for most people³⁰. Tonic heat pain has been successfully employed in the context of expectancy manipulations, with comparable temperatures^{31,32}. In previous studies, we have

	Placebo condition (PBO)	No-treatment condition (NT)
Instruction (either by VEx or Text DELIVERY)	"According to your randomization sheet, you have been assigned to the treatment condition . Please pick up the drug from the compartment and take it.[medication automatically dispensed] You will find a bottle of water there. Once you are finished, let me know."	"According to your randomization sheet, you have been assigned to the no treatment condition . This means that you will not receive any medication, but will simply repeat the hot water test."
Medication	Automatically dispensed, inert pill	None
Reinforcement 1	Questionnaire items addressing prior experience with the supposed medication, and its expected efficacy (expectation of pain relief), administered post treatment	n/a
Reinforcement 2	Questionnaire items supposedly assessing medication side effects, administered post treatment and post intermission	n/a

Table 1. Protocol differences between placebo and no-treatment EXPECTANCY. Beyond the core manipulation (inert medication versus no medication), instructions differed between groups, and the placebo group's expectancies were reinforced on two occasions.

established sufficient retest reliability to warrant the use of hot water in a baseline-corrected design³³. Due to the relatively high intensity, it was emphasized that they could discontinue the experiment at any time.

Subjects were seated comfortably in front of a two monitor-setup (instruction screen and touch screen for ratings), and the water bath. They underwent a familiarization trial where they briefly immersed their right hand. The actual experimental sessions consisted of immersions of the hand for 1 min, removal for 30 s, and reinsertion, for a total of 5 immersions each session. Immersion and removal were prompted by verbal displays on screen, and concurrent audio signals.

Throughout the sessions, subjects continually rated their pain on a 0–100 digital visual analogue scale (VAS) using a touch screen. However, we avoided using second-by-second data in our analyses because this would lead to overly complex models. Instead, we averaged continuous pain ratings to obtain mean pain scores for each of the 2 × 5 immersions.

After each session, using the touch screen, they also gave a single retrospective rating on a numerical rating scale from 0 to 10 for pain intensity and pain unpleasantness to establish comparability to conventional non-continuous measures, and rated their desire for pain relief. Furthermore, they responded to a tolerance item intended to gauge the extent to which they believe they could have continued the stimulation (Supplementary Table S1).

Experimental factor 1: instruction delivery. We created a Unity3D-based system for simulating the interactive text or virtual human (introduced as "Dr. Halsey") based system (Unity Technologies, San Francisco, CA, USA). The simulation was displayed on a large screen 65 in. display, where the virtual human was rendered as a life size character to be perceived as sitting right across the table from the subject. Additional detail for the development of the VEx is provided in Supplementary Text S2. A Microsoft Kinect motion tracking system (Microsoft, Redmond, WA) was installed above the display and tracked the subjects' head position in real-time. Using the head tracked position, Dr. Halsey's head orientation was animated such that Dr. Halsey would maintain socially realistic face-to-face gaze behaviors.

A finite state machine was created so that the simulation would trigger the logical step-by-step instructions in the VEx or Text DELIVERY conditions. Both conditions were displayed with the same duration. The simulation in the Text DELIVERY was designed such that at each step of the instruction, the written instructions on the screen were accompanied by the respective audio recording. The corresponding words on the screen were highlighted as they were spoken, to increase the similarity in attentional focus and pacing between Text and VEX DELIVERY. In the VEX DELIVERY, Dr. Halsey executed discrete instruction sets of verbal and non-verbal behaviors based on the same state machine, with animation transitions between the end of a state and the beginning of the next. In this manner, Dr. Halsey delivered the instruction via naturalistic verbal and non-verbal behaviors.

At interaction points during instruction delivery, subjects were required to provide an input to a question such as "yes" or "no", or select optional questions from an FAQ list, via a touchscreen interface. Either DELIVERY condition would provide appropriate responses to subject input before continuing with the step-by-step instructions. In the VEx condition when Dr. Halsey was awaiting input from the subjects, neutral life-like behaviors were displayed, such as breathing and small-scale head or hand movements.

Experimental factor 2: expectancy manipulation. The EXPECTANCY manipulation consisted of several steps intended to plausibly convey the notion that subjects have received an active treatment, when in fact, no active treatment was delivered. During instruction delivery, they were assigned to either the placebo (PBO) or the no-treatment control group (NT). Table 1 displays differences in EXPECTANCY manipulation between the two groups, most notably the "medication" delivered to PBO but not to NT. Of note, all other aspects of the instruction (including e.g. duration, information about the drug) were identical between PBO and NT.

The supposed medication (actually an inert lactose placebo) was delivered by a custom-made pill dispenser controlled by an Arduino microprocessor (<https://arduino.cc>), which was triggered simultaneously with the respective part of instruction delivery.

Psychological measures. Immediately after the intervention, subjects in the PBO EXPECTANCIES responded to items concerning the treatment, which also served as a co-intervention (see Table 1, Reinforcement 1 and Reinforcement 2). Reinforcement 1 was framed as medication-related with two corresponding items (“Do you have experience with Acetaminophen?” and “If you have experience with it, how effective was Acetaminophen in alleviating your symptoms?”). These served as distractors for a third item assessing the expectation of pain relief (“Considering the pain stimulus you received, how effective do you expect it to be in reducing the pain?”, with possible responses “Not effective at all”, “Not very effective”, “Fairly effective” and “Very effective”).

Throughout the experiment, we used several questionnaires to assess psychological traits and states; references are provided in Supplementary Table S2. Among the trait questionnaires were the Internality, Powerful Others and Chance Scales (IPC), the Revised Life Orientation Test assessing optimism and pessimism (LOT-R), the Big Five Inventory-44 assessing extraversion, openness to experience, conscientiousness, neuroticism and agreeableness (BFI-44), the trait subscale of the State-Trait Anxiety Inventory (STAI), as well as the Pain Vigilance and Awareness Questionnaire (PVAQ) and the Fear of Pain Questionnaire III (FPQ-III). The state questionnaires included the state subscale of the STAI, and the Multidimensional Mood Questionnaire assessing positive versus negative affect, wakefulness versus tiredness, and calm versus nervousness (MDMQ).

Most of the questionnaires have been associated with placebo responding before^{34–37}, but mood and BFI subscales were also intended for exploring affective influences on pain³⁸.

Furthermore, we obtained ratings on form and content of the treatment interaction to compare VEx and Text DELIVERY. These measures included several scales established in virtual reality research. In particular, we assessed social presence with the Social Presence Scale (SPS), and co-presence, attentional allocation, perceived message understanding and perceived behavioral interdependence with the Networked Minds Questionnaire (NMQ). Furthermore, we used a questionnaire specifically constructed for this experiment, which assessed the VEx/Text deliveries’ believability, likeability and the extent to which it was perceived as compelling (Expectancy Induction Characteristics scale, EIC).

Psychological measures were subjected to a quality assessment^{39,40} (see Supplementary Table S3). After individual assessment of all flagged questionnaires responses, none of the subjects displayed sufficient grounds for global exclusion.

Analysis. Due to the hierarchical nature of the data, we used a longitudinal hierarchical linear modeling (HLM) approach for analysis⁴¹. Among other advantages of HLM, imbalance in group sample sizes is unproblematic⁴². In the main HLM analyses, time variables (IMMERSION, SESSION) were entered as within-person variables. Second level (between-group) predictors (DELIVERY, EXPECTANCY) were then entered into the base model to ascertain the effects of the experimental conditions. All HLM analyses were performed with random effects including intercept, IMMERSION (linear and squared), and SESSION. All repeatedly measured predictors were separated into their within-person component, and between-group components (aggregated to the person level).

Secondary analyses were then performed for two purposes: to assess possible moderation of the main results by various second level variables (e.g. personality traits), and to see if other first level variables (e.g. retrospective ratings) were changed following the intervention. For the first purpose, covariates were separately entered into the main analysis. They included subject doubt, desire for pain relief, expectation of pain relief, and several psychological traits. For the second purpose, the dependent variable (pain) used in the main analysis was replaced by the respective other first level variables. They included retrospective ratings (pain intensity, pain unpleasantness, and pain tolerance) as well as mood and anxiety questionnaires (STAI, MDMQ). For all secondary analyses, alpha correction was performed within the respective domains using effective number of tests (M_{eff}) correction (Bonferroni correction adjusted for non-independence/correlation of predictors)⁴³. The domains were pain measures, pain-related psychological traits, personality-related psychological traits, expectation of medication effects, DELIVERY-related measures, psychological states.

Analyses were performed with SPSS 22 (SPSS Inc., Chicago, IL, USA) and in MATLAB R2017b (The MathWorks Inc., Natick, MA, USA), exploratory mediation analyses using MATLAB’s VBA toolbox⁴⁴. HLM analyses used Satterthwaite approximation for degrees of freedom. Significance level was set to 0.05. Descriptive statistics are reported as mean \pm SE unless otherwise noted.

Results

Baseline comparability and additional sampling issues. Fifty-four subjects were included in the study; the CONSORT flow diagram is provided in Fig. 3. Groups were comparable in all demographical and psychological state and trait measures assessed during baseline (see Supplementary Table S4). Completion of the psychological questionnaire battery took 16 ± 4 min (mean \pm SD).

Allocation to the 2×2 experimental groups was asymmetric such that both placebo EXPECTANCY conditions included more subjects than the control EXPECTANCY conditions (n of Text/PBO 17, VEx/PBO 17, Text/NT 11, VEx/NT 9). This was done to perform subgroup analyses excluding subjects with doubts about the cover story. Unfortunately, we were not able to solve the issue of inferring a priori doubt about the instructions (that is, doubt existing before retest) from our a posteriori measurement (manipulation check after retest). Regardless, results do not differ whether including putatively unconvinced subjects or not (data not shown).

Effects of immersion time on pain. The distribution of raw pain ratings is shown in Fig. 4. Interindividual differences were large, spanning almost the entire VAS. However, there are no conspicuous ceiling or floor effects.

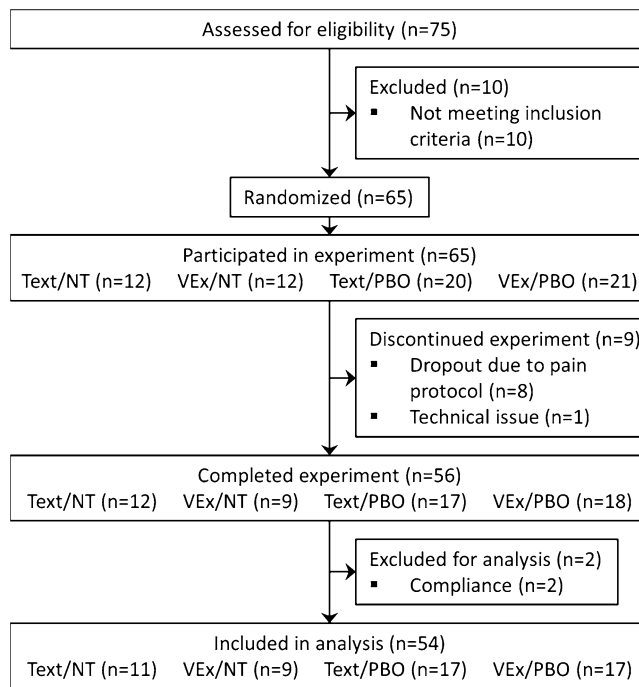


Figure 3. CONSORT statement for study enrollment. Attrition due to discontinued experiments was 14% (n = 9). Exclusions for compliance due to implausible core measures (pain ratings of zero) and antagonism (dozing off during baseline).

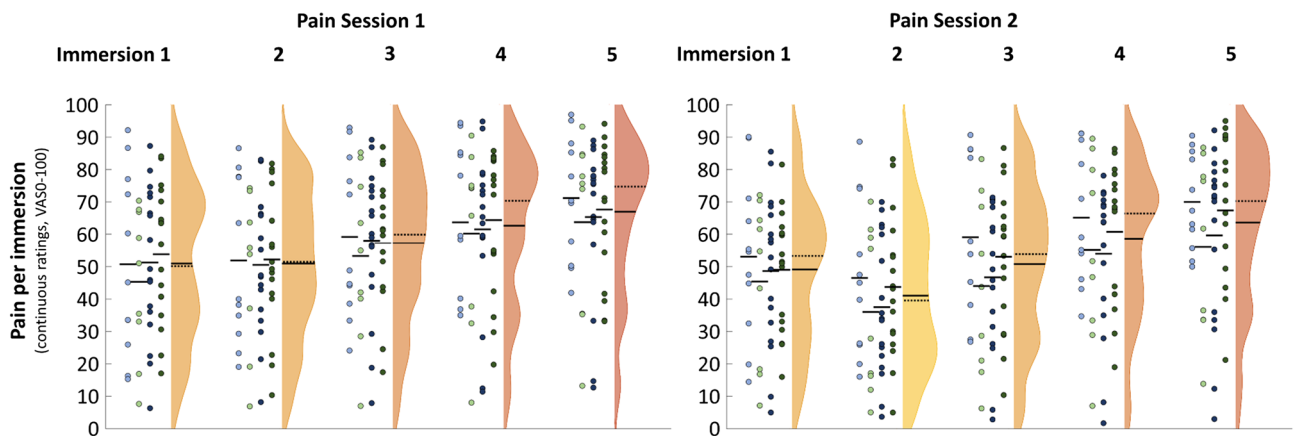


Figure 4. Pain ratings over both experimental sessions. Dot columns/colors indicate experimental groups (factor combinations of DELIVERY*EXPECTANCY; column 1 light blue Text/NT, column 2 light green VEx/NT, column 3 blue Text/PBO, column 4 green VEx/PBO). Violins show full sample distribution, mean (solid line) and median (dotted line) per immersion.

The hierarchical linear model indicated that within-session pain ratings follow a quadratic time course: higher ratings in immersion 1, equal or lower ratings in immersion 2, and a roughly linear increase between immersions 2 and 5 (see Supplementary Table S5 for full results). Furthermore, pain VAS decreases between sessions (-7.3 ± 1.5 VAS on average). Both findings are very robust and replicate our previous results³³.

Effect of experimental conditions. In addition to the base model (with first level-variables IMMERSION and SESSION), this model includes the effect of the two between-group experimental factors, EXPECTANCY and DELIVERY (Table 2, Supplementary Table S6). Importantly, we see an interaction of SESSION*EXPECTANCY*DELIVERY ($p=0.036$). Inspecting pain changes between sessions (Fig. 5), we see that the interaction arises from the difference of the two EXPECTANCY treatments in the Text DELIVERY (first column), that is, treatment differences arise according to our hypothesis (0.6 ± 3.1 VAS decrease in Text/NT, 8.0 ± 3.5 decrease VAS in Text/PBO; contrast $t(507) = -2.402$, $p=0.017$). Conversely, the VEX DELIVERY does not follow this intuitive mechanism, and

	Estimate	SE	CI lower	CI upper	p
SESSION*DELIVERY	1.126	3.027	- 4.822	7.073	0.711
SESSION*EXPECTANCY	- 3.857	3.135	- 10.016	2.302	0.222
IMMERSION*SESSION*DELIVERY	- 1.151	2.478	- 6.02	3.718	0.642
IMMERSION*SESSION*EXPECTANCY	- 0.765	2.571	- 5.816	4.286	0.766
SESSION*DELIVERY*EXPECTANCY	13.411	6.28	1.073	25.75	0.036

Table 2. Effects of immersion time and experimental conditions DELIVERY and EXPECTANCY on continuous pain VAS ratings (hierarchical linear model). For brevity, only effects are displayed that a) included either one of the experimental conditions, and the session term, and b) are of lower or equal order than three-way interactions (higher order all n.s. $p > 0.169$). For the full model, see Supplementary Table S6. *SE* standard error, *CI* 95% confidence interval.

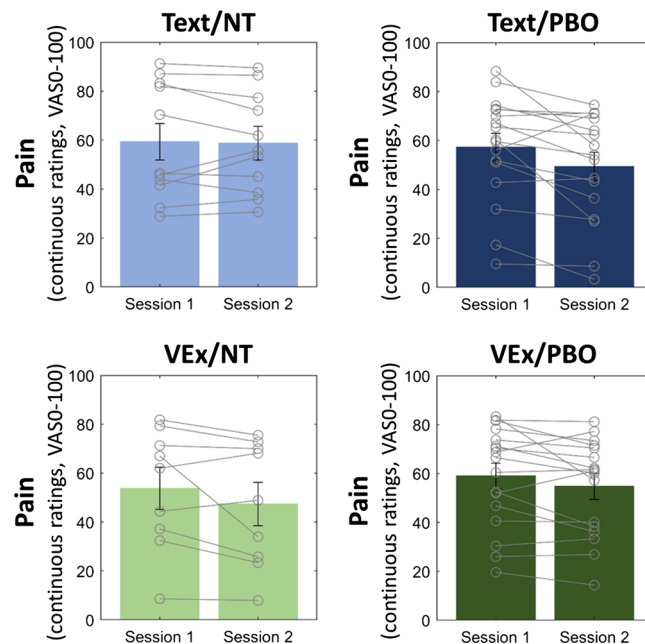


Figure 5. Change in mean pain ratings between session 1 and 2, by experimental condition (DELIVERY*EXPECTANCY).

there is no significant difference between the NT and PBO in this condition (6.4 ± 4.8 VAS decrease in VEx/NT, 4.2 ± 2.9 VAS decrease in VEx/PBO; contrast $t(507) = 0.676$, $p = 0.5$).

Secondary analyses

Moderation of expectancy manipulation by desire for pain relief and expectation of pain relief. Average values for desire for pain relief were high (mean \pm SD 6.13 ± 2.29 on a 0–10 scale; see Supplementary Table S4). Expectation of pain relief (0–3 scale) was assessed after the EXPECTANCY manipulation, and only in the PBO EXPECTANCIES, with no mean difference between the Text and VEx DELIVERY (mean \pm SD 1.41 ± 0.62 in Text, 1.59 ± 0.51 in VEx; two-sample t test $t(32) = -0.910$, $p = 0.370$).

Separately entering desire for pain relief (0–10 scale) and expectation of pain relief (0–3 scale) into the main analysis, a main effect of desire for pain relief indicated that a 1 point increase was associated with roughly 6.4 ± 1.1 higher VAS ratings over both sessions, across all conditions ($p < 0.0001$). Presumably, the causality here is that higher pain induces higher desire for its relief.

The expectation of pain relief significantly moderated pain change upon retest. Namely, we found a three-way interaction with SESSION and DELIVERY, such that in the Text DELIVERY, expectation had no effect on pain relief, but in the VEx DELIVERY, higher expectation preceded lower pain relief (Fig. 6) (SESSION*DELIVERY*EXPECTANCY $p = 0.027$).

Entering both desire for pain relief and expectation of pain relief simultaneously (for PBO only) did not change any of these results.

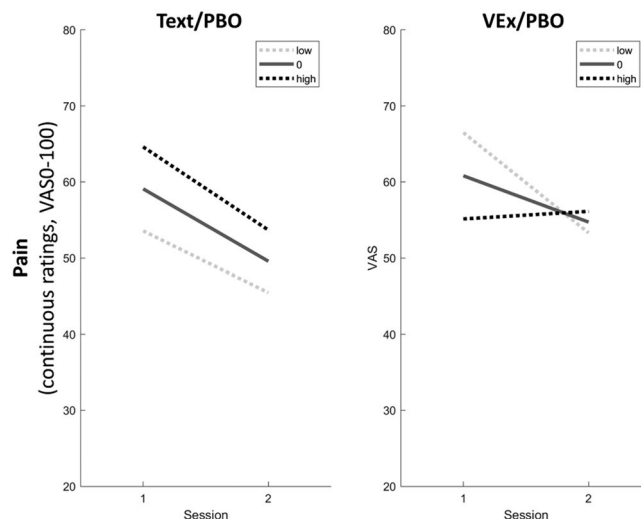


Figure 6. Moderation of the DELIVERY effect on pain by expectation of pain relief, controlling for desire for pain relief. In the Text DELIVERY (left panel), expectation has no effect (SESSION*Expectation $p=0.663$, single contrast; the only conspicuous effects are due to baseline differences), whereas in the VEx DELIVERY (right panel), higher expectation was associated with lower pain improvement at retest (SESSION*Expectation $p=0.004$, single contrast). Bright dotted line, low expectation of pain relief (-1 standard deviation from group means); solid line, average expectation of pain relief; dark dotted line, high expectation of pain relief ($+1$ standard deviation from group means).

Perception of the VEx compared to Text DELIVERY, and moderation by VEx perception. We found no significant correlations between either the hours spent on a computer or the hours spent using 3D technology, and any of the measures evaluating the VEx such as co-presence [highest $r(24)=0.28$, $p(\text{uncorr.})=0.172$].

Individually entering each of the VEx measures as dependent variables in fixed effects models including DELIVERY and EXPECTANCY as predictors, we found that DELIVERY had a significant main effect on co-presence after correction for $M_{\text{eff}}=7.094$ [6.46 ± 1.64 higher NMQ in VEx DELIVERY; $F(1,50)=14.33$, $p(\text{corr.})=0.003$]. This is unsurprising, as the co-presence scale contains items that strongly imply the presence of an embodied agent, and the visual perception of its attending the observer (e.g. “Dr. Halsey noticed me”, “I caught Dr. Halsey’s attention”). No other variable was significantly influenced by either DELIVERY or EXPECTANCY.

Individually entering each of the VEx measures as moderators in the main hierarchical linear model did not yield any significant findings after correction for multiple comparisons (Supplementary Table S7).

Moderation of expectancy manipulation by three pain-related subscales. Three pain-related subscales were entered into the main analysis individually, namely the PVAQ and the minor and severe pain subscales of the FPQ (Supplementary Table S8). We found no significant associations with or without correction for $M_{\text{eff}}=2.97$ tests.

Moderation of expectancy manipulation by ten personality-related subscales. Next, we individually entered the ten personality-related subscales from IPC, LOT-R, BFI and STAI into the main analysis. These analyses did not yield any significant findings after correction for multiple comparisons for $M_{\text{eff}}=9.162$ [lowest $p(\text{corr.})=0.054$]. For exploratory reasons, we want to point out that trait optimism (LOT-R), the internality subscale of the IPC and the conscientiousness subscale of the BFI-44 show a moderation effect at uncorrected significance level (see Supplementary Table S9).

Effect of experimental conditions on retrospective ratings. Next, we used the three retrospective ratings (0–10 numerical rating scales) individually as dependent variables, corrected for $M_{\text{eff}}=2.061$ tests. For pain intensity, this yielded a significant effect of the experimental conditions (SESSION*EXPECTANCY interaction, $p(\text{corr.})=0.023$). Single contrasts indicate that this effect was mostly driven by the difference between Text/NT (which actually reported increasing pain intensity between sessions) and the two PBO conditions (which reported decreasing pain intensity between sessions) (Fig. 7a). For the pain tolerance measure, a similar effect was found [SESSION*EXPECTANCY interaction, $p(\text{corr.})=0.016$; Fig. 7b]. Note that model residuals did not show a deviation from normality (Shapiro–Wilk test $p=0.694$), so we deemed the test as appropriate for the Guttman scale used for pain tolerance. We found no effect of experimental conditions for pain unpleasantness.

Effect of experimental conditions on psychological state variables. Four psychological state variables were entered as alternative dependent variables to assess the effect of the experimental conditions, corrected for $M_{\text{eff}}=2.598$ tests. For the state anxiety scale of the STAI, contrary to the retrospective pain ratings where a clear main effect of PBO became visible (that is, significant SESSION*EXPECTANCY interaction), an

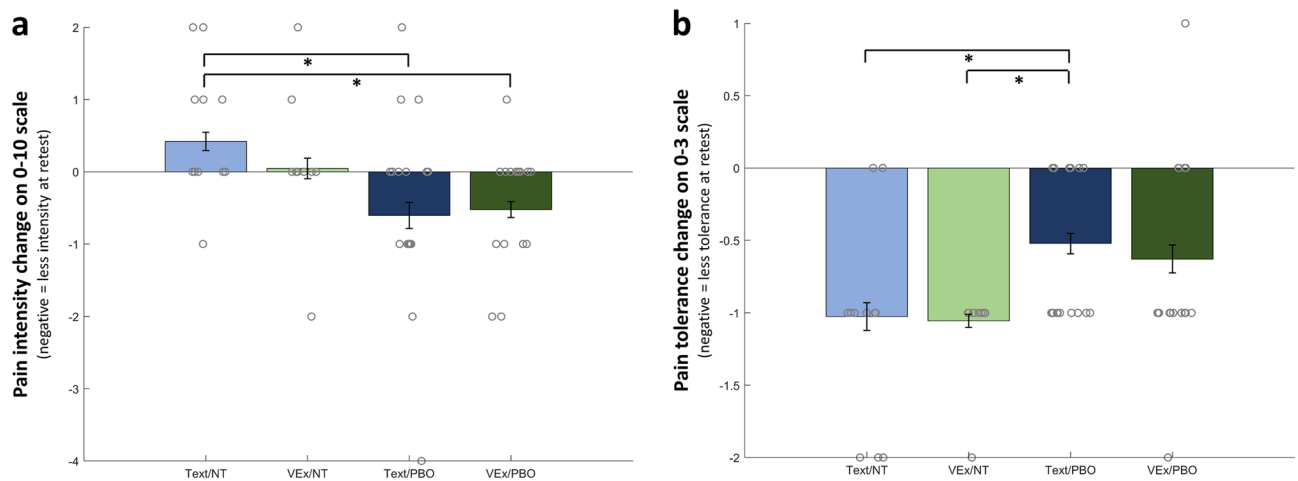


Figure 7. Changes in retrospective ratings between baseline and retest for the four groups. Grey circles are single data points (score differences). Legend: *, $p < 0.05$. (a) Retrospective pain intensity ratings. (b) Retrospective pain tolerance ratings.

interaction of SESSION*DELIVERY*EXPECTANCY reached significance [$p(\text{corr.}) = 0.034$; Fig. 8a]. This effect arose because in the Text/NT condition, a large increase in anxiety (~8 pts in 20–80 scale) occurred between the sessions which was absent in VEx/NT (single contrast $p = 0.022$) and Text/PBO (single contrast $p = 0.022$); however, no difference in anxiety changes were seen compared to the VEx/PBO condition ($p = 0.149$). Both the MDMQ Good/Bad scale and the MDMQ Calm/Nervous scale complement this finding, with three-way interactions on Good/Bad [$p(\text{corr.}) = 0.017$; Fig. 8b] and Calm/Nervous [$p(\text{corr.}) = 0.015$; Fig. 8c], driven by differences of Text/NT to VEx/NT and Text/PBO but not VEx/PBO.

As the effect of the experimental conditions on the three state variables closely mirrored those on continuous pain, we performed mediation analyses to see whether the pain changes were in any way related to the changes in anxiety. To do so, we entered difference scores of continuous pain between sessions as dependent variable, difference scores between the state variables as mediators, and performed separate analyses for each contrast of the experimental conditions. None of the analyses indicated significant mediation of pain change by any of the state variables (all $p > 0.1$ for Sobel tests of mediation, with $p = 0.106$ for mediation through the MDMQ Calm/Nervous subscale).

Discussion

To investigate whether the delivery of treatment instructions could be fully standardized, while still preserving the necessary social interactions required for the generation of expectancy effects, we tested a virtual experimenter (VEx) against a text-only condition in a baseline–retest placebo manipulation. Overall, the automated delivery of a treatment was followed by substantial pain improvement. This is promising, because it means that some sources of bias in placebo research can be eliminated, such as bias from unblinding or lack of standardization.

Contrary to our hypotheses, we found no larger pain reduction when instructions were delivered by an embodied virtual agent, compared to a text-only condition. The main result was found in the Text DELIVERY with significant differences between PBO and NT group, such that pain decreased after placebo manipulation but not after no-treatment. However, this effect was absent in the VEx DELIVERY. Instead, placebo and no-treatment group both improved somewhat, but the effect was driven by a larger pain decrease in the no-treatment condition. This finding is intriguing, as it suggests that differences in the formal aspects of communication (embodied agent versus text), with identical explicit information content, could lead to effects as strong as “classical” expectancy effects observed under text-only, placebo versus no-treatment conditions.

Disappointingly, these results indicate that, in the current protocol, we have to reject our hypothesis of improved placebo responding after interacting with the VEx. This also implies that the intended “homogenization” of placebo effects through increased standardization failed to manifest. We note that the VEx was well-accepted but had no obvious superiority in terms of VR-related measures, as compared to the Text condition, beyond the co-presence measure attesting to the existence of an embodied agent. In other words, within the limited interaction employed in this experiment, measures such as attentional focus and message understanding did not benefit from the embodiment of the agent. Neither the measures concerning the VEx itself, nor its correlates in psychological state variables, offered straightforward explanations for these findings. It seems unlikely that the embodiment abolishes placebo responding altogether, given the effects on retrospective ratings discussed below, which directly compare to Text DELIVERY. Granted the between-group design with comparatively large variances, it is likely that additional research is needed, possibly with larger sample sizes.

On state anxiety, nervousness and mood, the experimental conditions showed analogous effects to pain rating scales. It has been suggested that expectancy effects on pain partially work by affecting anxiety or stress^{45–47}. In our case however, the absence of mediation effects suggests that either the effect on continuous pain and on psychological states were not observed in the same subjects, or that the association was too small for detection. Indeed, the study was not powered to conclusively perform this type of analysis. In either case, it is likely

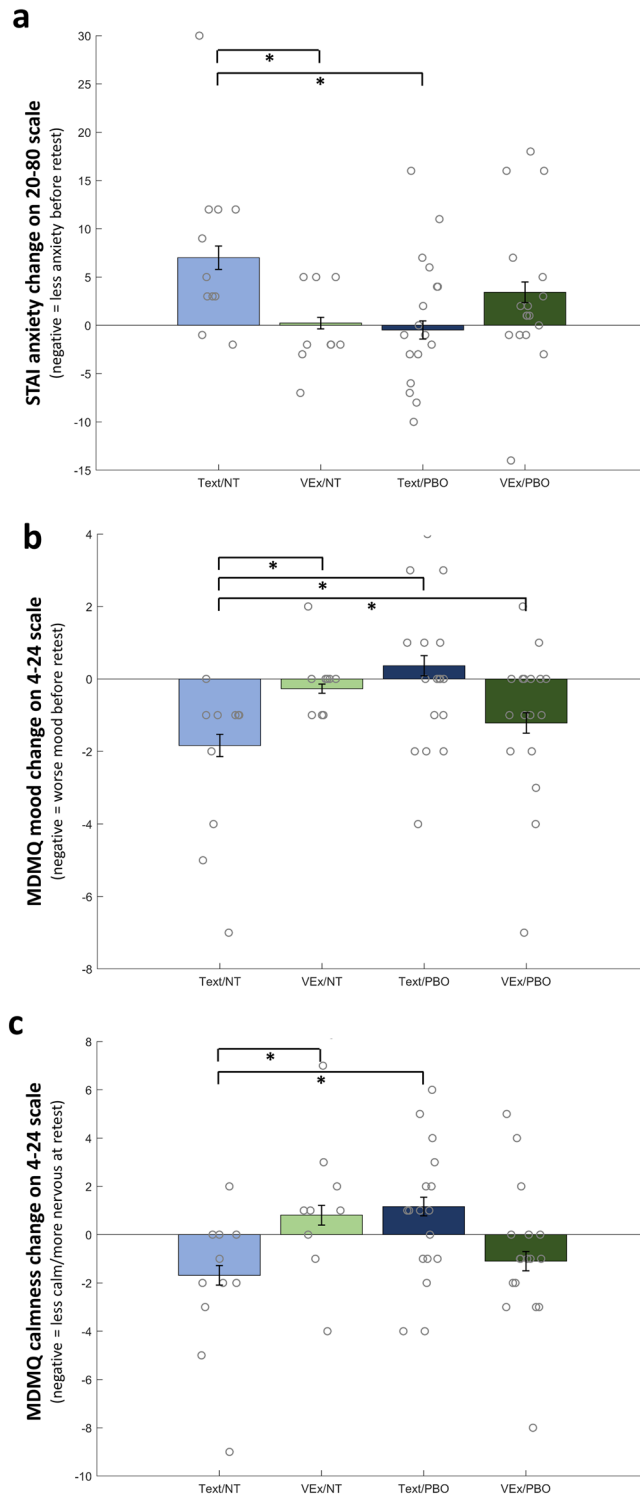


Figure 8. Changes in psychological states between baseline and retest for the four groups, measured prior to the painful stimulation. Grey circles are single data points (score differences). **(a)** Change in STAI anxiety. **(b)** Change in MDMQ Good/Bad. **(c)** Change in MDMQ Calm/Nervous.

that complex psychological interactions are at the core of this finding, which is not exhaustively explained by expectancy.

For example, high expectation of pain relief, as measured in the PBO conditions only, had no effect in the Text DELIVERY, but led to a *lesser* pain decrease in the VEx DELIVERY. This is contrary to intuition and previously reported results^{48,49}. Salient expectancy violations may have led to a contrast effect—instead of integrating the expectations of symptom relief (“assimilation”), a paradoxical symptom worsening occurs—or more generally,

an aversive affective response⁵⁰. Research documenting such effects exists on both the behavioral and neuronal level^{51–53}. However, the identification of mechanisms is hindered by the fact that such effects do not reliably manifest. Therefore, the precise (experimental, much less clinical) conditions for when assimilation turns to a contrast effect are a matter of speculation⁵⁴, for example, involving differences in sensory precision, attention, trust or prior experiences. Fittingly, if expectancy violations were to explain the effect of expectation of pain relief in the VEx group, they would fail to do so in the Text group where the PBO EXPECTANCY was assimilated.

Alternatively, while no corresponding data exists, we speculate that higher explicit expectations may have entailed a higher allocation of attentional resources to the pain experienced at retest. This would in turn make perception of the immediate sensory input more precise. As has been suggested⁵⁵ and supported using a heat pain protocol⁵⁶, when placebo hypoalgesia is conceptualized as a form of Bayesian integration, higher precision of sensory input would actually reduce placebo effects by reducing the impact of the expectancy prior.

Either interpretation does not explain the finding of comparatively large pain reduction in the VEx/NT condition. As we have argued before^{34,57}, in control conditions, other factors such as a high internal locus of control may account for the extent to which aversive symptoms are experienced if only personal coping resources are available (i.e., no external attribution is possible, like on received medication). Our data somewhat corroborates this notion with a corresponding effect in the VEx/NT prior to alpha correction (higher internality leading to higher pain reduction), although internality had a parallel effect in the Text/PBO condition.

In any case, findings in the VEx condition are unlikely to be explained by an effect of distraction. While distraction can be a highly effective strategy to reduce pain⁵⁸, the VEx manipulation was separated from the pain retest by a 30 min waiting period and not displayed concurrently to painful stimulation (as in⁵⁹, for example). At best, a reminiscence of novelty may have persisted.

Furthermore, it is noteworthy that the described effects are only visible in continuous pain or indirect measures such as mood and anxiety. In the majority of studies of expectancy mechanisms, continuous pain measurements are not available, mostly due to more transient pain stimuli, and for methodological concerns that concurrent ratings may in itself be distracting. Instead, ratings are usually assessed after the respective pain trials. In our case, such retrospective ratings do not show the interaction between EXPECTANCY and DELIVERY conditions, but flat effects of EXPECTANCY on SESSION differences, across DELIVERY conditions. A number of studies have pointed out regularities in the way pain is remembered (for an overview, see⁶⁰). Most notably, this includes the peak-end rule^{61,62}, which states that people tend to remember experiences mostly based on how they were perceived at their most intense, and at the end of the experience. In our study, the peak almost invariably coincides with the end. Therefore, one could argue that the continuous ratings including the entire previous pain trajectory merely introduce noise irrelevant for pain memory. Both pain experience and pain memory are important, one being a primary measure to reduce suffering, the other paramount for shaping future behavior, which is a crucial function of pain⁶³. The latter function could involve, for example, dysfunctional behaviors implied by the fear-avoidance model⁶⁴, or (non)adherence to treatments⁶². An alternative interpretation of the disconnect between continuous and retrospective ratings may be that continuous ratings are less amenable to demand characteristics or similar response biases because of high working memory demands to provide them. Likewise, mood and anxiety are less transparently related to the expectancy manipulation and therefore may be less influenced by it. More dedicated studies are needed to address these issues.

We were satisfied to see that the pain protocol had sufficient variance between subjects that we could detect a significant effect of the EXPECTANCY manipulation. We also opted against a within-person, cross-over design routinely employed in placebo research, which might have further reduced error variance, out of concerns for asymmetric carry-over effects³⁴. The existence of a significant EXPECTANCY effect is all the more noteworthy, as the simple expectancy manipulation we used (verbal suggestion of efficacy) is the most basic and arguably weakest form of manipulation employed in placebo mechanism studies^{65,66}. While we have reinforced the placebo instruction intermittently (see Table 1), several other options exist to strengthen this protocol aspect now that basic efficacy was established⁶⁷, the most obvious being to include some form of conditioning procedure (e.g.^{49,66}). This is particularly important given that confidence intervals were comparatively large.

As an additional limitation, the study was not powered to investigate higher-level interactions of EXPECTANCY and other determinants. For example, complex interactions between EXPECTANCY, sex, optimism and anxiety may be involved in the generation of placebo effects⁶⁸. The observed interactions with our experimental conditions support the notion that the identification of “simple” predictors of placebo responding³⁴ may be a fool’s errand, with more interactional processes likely but harder to investigate^{35,69}. Intriguingly, similarities and differences found between our experimental groups may eventually be explained by the way expectancies are formed in the first place, by different methods and in different people^{5,70}.

In conclusion, we have demonstrated a fully standardized and automated procedure capable to deliver circumscribed treatments and hence generate expectancy effects. These responses did not scale with closer semblance to a human experimenter, as DELIVERY and EXPECTANCY did not exhibit a synergistic effect as hypothesized. Instead, both had comparable effect size but seem to rely on different mechanisms. Putatively, a higher salience of the manipulation in the VEx DELIVERY was a driving factor for this divergence. Effects were moderated by expectation of pain relief and, on average, mimicked by those on psychological state variables.

Future directions should include replication in a more robust placebo protocol and/or a larger sample size. The addition of an actual human experimenter as a third mode of DELIVERY, originally intended¹⁸ but scrapped for logistical reasons, should be strongly considered to investigate psychosocial determinants of treatment success. The interaction with the virtual agent could be expanded such that the VEx is also “present” during the pain sessions themselves. It would be very promising to eventually transfer this paradigm into an actual clinical trial to see how the effects of the treatment compare to its mode of delivery, with the ultimate aim of utilizing similar technology in clinical care itself⁷¹. Future manipulations could utilize internet-based and therefore highly scalable data collection such as Pavlovia (www.pavlovia.org) or Amazon Mechanical Turk (www.mturk.com). In its

current state, systematic examination of overt traits such as age, sex, race or (non-)verbal behaviors in various healthy and patient populations could be an avenue to further the understanding of the mechanisms of placebo effects⁷², and how they affect ubiquitous symptoms like pain. Even in a purely experimental context, a strength of the VEx is that it allows for a tightly controlled manipulation of such overt characteristics of appearance and behavior⁷³, and offers promise to the mitigation of bias in clinical research protocols.

Received: 4 August 2020; Accepted: 10 November 2020

Published online: 23 November 2020

References

- Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 743–800 (2015).
- Manchikanti, L., Singh, V., Kaye, A. D. & Hirsch, J. A. Lessons for better pain management in the future: Learning from the past. *Pain Ther.* <https://doi.org/10.1007/s40122-020-00170-8> (2020).
- Vase, L., Riley, J. L. & Price, D. D. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain* **99**, 443–452 (2002).
- Peerdeman, K. J. *et al.* Relieving patients' pain with expectation interventions. *Pain* **157**, 1179–1191 (2016).
- Vase, L. & Wartolowska, K. Pain, placebo, and test of treatment efficacy: A narrative review. *Br. J. Anaesth.* **123**, e254ee262 (2019).
- Enck, P., Bingel, U., Schedlowski, M. & Rief, W. The placebo response in medicine: Minimize, maximize or personalize?. *Nat. Rev. Drug Discov.* **12**, 191–204 (2013).
- Benedetti, F., Carlino, E. & Pollo, A. Hidden administration of drugs. *Clin. Pharmacol. Ther.* **90**, 651–661 (2011).
- Colloca, L. & Grillon, C. Understanding placebo and nocebo responses for pain management. *Curr. Pain Headache Rep.* **18**, 419 (2014).
- Kelley, J. M. *et al.* Patient and practitioner influences on the placebo effect in Irritable Bowel Syndrome. *Psychosom. Med.* **71**, 789–797 (2009).
- Bingel, U. *et al.* The effect of treatment expectation on drug efficacy: Imaging the analgesic benefit of the opioid remifentanyl. *Sci. Transl. Med.* **3**, 70ra14 (2011).
- Petrie, K. J. & Rief, W. Psychobiological mechanisms of placebo and nocebo effects: Pathways to improve treatments and reduce side effects. *Annu. Rev. Psychol.* **70**, 599–625 (2019).
- Ioannidis, J. P. A. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* **294**, 218 (2005).
- Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* (80–) **349**, aac4716 (2015).
- Hróbjartsson, A., Forfang, E., Haahr, M., Als-Nielsen, B. & Brorson, S. Blinded trials taken to the test: An analysis of randomized clinical trials that report tests for the success of blinding. *Int. J. Epidemiol.* **36**, 654–663 (2007).
- Schwarz, K. A. *et al.* How stereotypes affect pain. *Sci. Rep.* **9**, 8626 (2019).
- Rosenthal, R. & Jacobson, L. *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development* (Holt, Rinehart and Winston, New York, 1968).
- Hróbjartsson, A., Kaptchuk, T. J. & Miller, F. G. Placebo effect studies are susceptible to response bias and to other types of biases. *J. Clin. Epidemiol.* **64**, 1223–1229 (2011).
- Horing, B., Newsome, N. D., Enck, P., Babu, S. V. & Muth, E. R. A virtual experimenter to increase standardization for the investigation of placebo effects. *BMC Med. Res. Methodol.* **16**, 84 (2016).
- Carl, E. *et al.* Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *J. Anxiety Disord.* **61**, 27–36 (2019).
- Solcà, M. *et al.* Heartbeat-enhanced immersive virtual reality to treat complex regional pain syndrome. *Neurology* **91**, e479–e489 (2018).
- Triberti, S., Repetto, C. & Riva, G. Psychological factors influencing the effectiveness of virtual reality-based analgesia: A systematic review. *Cyberpsychol. Behav. Soc. Netw.* **17**, 335–345 (2014).
- Colloca, L. *et al.* Virtual reality: Physiological and behavioral mechanisms to increase individual pain tolerance limits. *Pain* <https://doi.org/10.1097/j.pain.0000000000001900> (2020).
- Dahlquist, L. M. *et al.* Effects of videogame distraction using a virtual reality type head-mounted display helmet on cold pressor pain in children. *J. Pediatr. Psychol.* **34**, 574–584 (2009).
- Daugherty, B., Van Wallendael, L., Babu, S., Cutler, B. & Hodges, L. F. Virtual human versus human administration of photographic lineups. *IEEE Comput. Graph. Appl.* **28**, 65–75 (2008).
- Cutler, B. L., Daugherty, B., Babu, S., Hodges, L. & Van Wallendael, L. Creating blind photoarrays using virtual human technology. *Police Q.* **12**, 289–300 (2009).
- Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom. Med.* **64**, 258–266 (2002).
- Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The PHQ-9: Validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**, 606–613 (2001).
- Horn, C., Schaller, J. & Lautenbacher, S. Investigating the affective component of pain: No startle modulation by tonic heat pain in startle responsive individuals. *Int. J. Psychophysiol.* **84**, 254–259 (2012).
- Rainville, P. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* (80–) **277**, 968–971 (1997).
- Magerl, W. *et al.* Reference data for quantitative sensory testing (QST): Refined stratification for age and a novel method for statistical comparison of group data. *Pain* **151**, 598–605 (2010).
- Matre, D. Placebo-induced changes in spinal cord pain processing. *J. Neurosci.* **26**, 559–563 (2006).
- Lyby, P. S., Aslaksen, P. M. & Flaten, M. A. Is fear of pain related to placebo analgesia?. *J. Psychosom. Res.* **68**, 369–377 (2010).
- Horing, B., McCubbin, J. A., Moore, D. & Muth, E. R. Resting blood pressure differentially predicts time course in a tonic pain experiment. *Psychophysiology* **53**, 1600–1607 (2016).
- Horing, B., Weimer, K., Muth, E. R. & Enck, P. Prediction of placebo responses: A systematic review of the literature. *Front. Psychol.* **5**, 1–10 (2014).
- Geers, A. L., Kosbab, K., Helfer, S. G., Weiland, P. E. & Wellman, J. A. Further evidence for individual differences in placebo responding: An interactionist perspective. *J. Psychosom. Res.* **62**, 563–570 (2007).
- Morton, D. L., Watson, A., El-Dereby, W. & Jones, A. K. P. Reproducibility of placebo analgesia: Effect of dispositional optimism. *Pain* **146**, 194–198 (2009).
- Vachon-Presseau, E. *et al.* Brain and psychological determinants of placebo pill response in chronic pain patients. *Nat. Commun.* **9**, 3397 (2018).
- Lumley, M. A. *et al.* Pain and emotion: A biopsychosocial review of recent research. *J. Clin. Psychol.* **67**, 942–968 (2012).

39. Meade, A. W. & Craig, S. B. Identifying careless responses in survey data. *Psychol. Methods* **17**, 437–455 (2012).
40. DeSimone, J. A., Harms, P. D. & DeSimone, A. J. Best practice recommendations for data screening. *J. Organ. Behav.* **36**, 171–181 (2015).
41. Raudenbush, S. W. & Bryk, A. S. *Hierarchical Linear Models: Applications and Data Analysis* (Sage, Thousand Oaks, 2002).
42. Maas, C. J. M. & Hox, J. J. Sufficient sample sizes for multilevel modeling. *Methodology* **1**, 86–92 (2005).
43. Nyholt, D. R. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–769 (2004).
44. Daunizeau, J., Adam, V. & Rigoux, L. VBA: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput. Biol.* **10**, e1003441 (2014).
45. Flaten, M. A., Aslaksen, P. M., Lyby, P. S. & Bjørkedal, E. The relation of emotions to placebo responses. *Philos. Trans. R. Soc. B Biol. Sci.* **366**, 1818–1827 (2011).
46. Elsenbruch, S., Roderigo, T., Enck, P. & Benson, S. Can a brief relaxation exercise modulate placebo or nocebo effects in a visceral pain model?. *Front. Psychiatry* **10**, 20 (2019).
47. Aslaksen, P. M., Zwarg, M. L., Eilertsen, H.-I.H., Gorecka, M. M. & Bjørkedal, E. Opposite effects of the same drug. *Pain* **156**, 39–46 (2015).
48. Koyama, T., McHaffie, J. G., Laurienti, P. J. & Coghill, R. C. The subjective experience of pain: Where expectations become reality. *Proc. Natl. Acad. Sci.* **102**, 12950–12955 (2005).
49. Corsi, N. & Colloca, L. Placebo and nocebo effects: The advantage of measuring expectations and psychological factors. *Front. Psychol.* **8**, 20 (2017).
50. Geers, A. L. & Lassiter, G. D. Affective expectations and information gain: Evidence for assimilation and contrast effects in affective experience. *J. Exp. Soc. Psychol.* **35**, 394–413 (1999).
51. Colloca, L., Schenk, L. A., Nathan, D. E., Robinson, O. & Grillon, C. When therapeutic expectancies are violated: An fMRI study. *Clin. Pharmacol. Ther.* **106**, 1246–1252 (2019).
52. Zeidan, F., Lobanov, O. V., Kraft, R. A. & Coghill, R. C. Brain mechanisms supporting violated expectations of pain. *Pain* **156**, 1772–1785 (2015).
53. Hird, E. J., Charalambous, C., El-Derehy, W., Jones, A. K. & Talmi, D. Boundary effects of expectation in human pain perception. *Sci. Rep.* **9**, 1–13 (2019).
54. Peerdeman, K., Geers, A., Porta, D. Della, Veldhuijzen, D. S. & Kirsch, I. Underpredicting pain: An experimental investigation into the benefits and risks—Preprint. <https://doi.org/10.31234/OSF.IO/YR9JT>.
55. Büchel, C., Geuter, S., Sprenger, C. & Eippert, F. Placebo analgesia: A predictive coding perspective. *Neuron* **81**, 1223–1239 (2014).
56. Grahl, A., Onat, S. & Büchel, C. The periaqueductal gray and Bayesian integration in placebo analgesia. *Elife* **7**, e32930 (2018).
57. Horing, B., Weimer, K., Muth, E. R. & Enck, P. Prediction of symptom change in placebo versus no-treatment group in experimentally induced motion sickness. *Appl. Psychophysiol. Biofeedback* **40**, 163–172 (2015).
58. Sprenger, C. *et al.* Attention modulates spinal cord responses to pain. *Curr. Biol.* **22**, 1019–1022 (2012).
59. Gupta, A., Scott, K. & Dukewich, M. Innovative technology using virtual reality in the treatment of pain: Does it reduce pain via distraction, or is there more to it?. *Pain Med.* **19**, 151–159 (2018).
60. Adamczyk, W. M. *et al.* Memory of pain in adults: A protocol for systematic review and meta-analysis. *Syst. Rev.* **8**, 201 (2019).
61. Kahneman, D., Fredrickson, B. L., Schreiber, C. A. & Redelmeier, D. A. When more pain is preferred to less: Adding a better end. *Psychol. Sci.* **4**, 401–405 (1993).
62. Redelmeier, D. A. & Kahneman, D. Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain* **66**, 3–8 (1996).
63. Seymour, B. Pain: A precision signal for reinforcement learning and control. *Neuron* **101**, 1029–1041 (2019).
64. Vlaeyen, J. W. S., Kole-Snijders, A. M. J., Boeren, R. G. B. & van Eek, H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* **62**, 363–372 (1995).
65. Wolters, F., Peerdeman, K. J. & Evers, A. W. M. Placebo and nocebo effects across symptoms: From pain to fatigue, dyspnea, nausea, and itch. *Front. Psychiatry* **10**, 470 (2019).
66. Rhudy, J. L., Güereca, Y. M., Kuhn, B. L., Palit, S. & Flaten, M. A. The influence of placebo analgesia manipulations on pain report, the nociceptive flexion reflex, and autonomic responses to pain. *J. Pain* **19**, 1257–1274 (2018).
67. Bishop, F. L. *et al.* What techniques might be used to harness placebo effects in non-malignant pain? A literature review and survey to develop a taxonomy. *BMJ Open* **7**, e015516 (2017).
68. Zhou, L. *et al.* The influence of expectancy level and personal characteristics on placebo effects: Psychological underpinnings. *Front. Psychiatry* **10**, 20 (2019).
69. Fava, G. A., Guidi, J., Rafanelli, C. & Rickels, K. The clinical inadequacy of the placebo model and the development of an alternative conceptual framework. *Psychother. Psychosom.* **86**, 332–340 (2017).
70. Geers, A. L., Briñol, P. & Petty, R. E. An analysis of the basic processes of formation and change of placebo expectations. *Rev. Gen. Psychol.* **23**, 211–229 (2019).
71. Oran, D. P. & Topol, E. J. The rise of the virtualist. *Lancet* **394**, 17 (2019).
72. Daniali, H. & Flaten, M. A. A qualitative systematic review of effects of provider characteristics and nonverbal behavior on pain, and placebo and nocebo effects. *Front. Psychiatry* **10**, 242 (2019).
73. Bernstein, M. H. *et al.* Putting the 'art' into the 'art of medicine': The under-explored role of artifacts in placebo studies. *Front. Psychol.* **11**, 1354 (2020).

Acknowledgements

We are grateful to Joseph P. Mazer, Nathan D. Newsome and Matias Volonte for their contributions in prototyping and development of the virtual experimenter, and to Paul Enck for valuable advice concerning both research and manuscript. This research has been supported by Clemson University's Human Factors Institute and the Alexander von Humboldt-Foundation (Feodor Lynen Fellowship).

Author contributions

B.H., E.R.M. and S.V.B. conceived and planned the experiment. S.C.B. carried out the experiment. Z.I. created the virtual reality simulation with supervision from A.R. and S.V.B. B.H. analyzed the data. B.H., E.R.M., A.R. and S.V.B. contributed to the interpretation of the results. B.H. wrote the manuscript with contributions from Z.I. and S.V.B. B.H. created the figures. All authors provided critical feedback and helped shape the research, analysis and manuscript. All authors reviewed the manuscripts.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-77453-9>.

Correspondence and requests for materials should be addressed to B.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020