Classical Conditioning Fails to Elicit Allodynia in an Experimental Study with Healthy Humans

Pain Medicine 2017; 18: 1314–1325
doi: 10.1093/pm/pnw221

OXFORD

METHODOLOGY, MECHANISMS & TRANSLATIONAL RESEARCH SECTION

Original Research Article

Classical Conditioning Fails to Elicit Allodynia in an Experimental Study with Healthy Humans

Victoria J. Madden, PhD,* Leslie N. Russek, PhD,[†,‡] Daniel S. Harvie, PhD,[*,‖] Johan W.S. Vlaeyen, PhD,[§,¶] and G. Lorimer Moseley, PhD[*,‖]

*Sansom Institute for Health Research; [†]University of South Australia, Adelaide, South Australia; [‡]Department of Physical Therapy, Clarkson University, Potsdam, New York, USA; [§]Research Group Health Psychology, KU Leuven, Leuven, Belgium; [¶]Department Clinical Psychological Science, Maastricht University, Maastricht, Netherlands; [‖]Neuroscience Research Australia, Sydney, Australia; [‖]Recover Injury Research Centre, Menzies Health Institute Queensland, Griffith University.

Correspondence to: G. Lorimer Moseley, Body in Mind Research Group, GPO Box 2471, University of South Australia, Adelaide 5001, Australia.
Tel: +61-8-8302-1416;
E-mail: lorimer.moseley@gmail.com.

Funding sources: VJM was supported by the Oppenheimer Memorial Trust, South Africa, and is now supported by an Innovation Postdoctoral Scholarship from the National Research Foundation of South Africa. JWSV is supported by the Odysseus Grant, ''The Psychology of Pain and Disability Research Program,'' funded by the Research Foundation Flanders (FWO-Vlaanderen), Belgium, as well as by the "Asthenes" long-term structural funding–Methusalem grant by the Flemish Government, Belgium. GLM is supported by an Australian National Health & Medical Research Council Principal Research Fellowship—"The Role of the Brain and Mind in Chronic Pain" (ID 1061279). This project supported by a project grant from the Australian National Health & Medical Research Council (ID 1047317). GLM has received support from Pfizer, Agile Physiotherapy, Results Physiotherapy, Kaiser Permanente, Return to Work South Australia. He receives Speaker fees for lectures on pain and rehabilitation and royalties for books on pain and rehabilitation.

Disclosure and conflicts of interest: All authors declare no conflicts of interest relevant to this work.

Abstract

Objective. Associative learning has been proposed as a mechanism behind the persistence of pain after tissue healing. The simultaneous occurrence of nociceptive and non-nociceptive input during acute injury mimics the pairings thought to drive classical conditioning effects. However, empirical evidence for classically conditioned allodynia is lacking. We aimed to manipulate pain thresholds with a classical conditioning procedure that used non-nociceptive somatosensory stimuli as conditioned stimuli (CS) and nociceptive stimuli as unconditioned stimuli. We also explored the influence of gender, depression, anxiety, negative affect, and pain catastrophizing on the main manipulation.

Design. Thirty-four healthy humans participated in a differential classical conditioning procedure that used vibrotactile stimulations at two different locations as CS. In an acquisition phase, CS+ was paired with painful thermal stimulation, and CS- with nonpainful thermal stimulation. Heat pain threshold was assessed during paired heat-CS trials before and after acquisition. A 2 (time: 1 and 2) x 2 (condition: CS+ and CS-) repeated-measures analysis of variance compared pain thresholds before and after acquisition. Exploratory analyses explored the influence of gender, depression,

anxiety, negative affect, and pain catastrophizing. Postexperiment questions investigated participants' awareness of the contingencies employed.

Results. The classical conditioning procedure did not alter pain thresholds. Exploratory analyses did not reveal any influence of individual differences. Thirty of the 34 participants were unaware of the contingencies between stimuli.

Conclusions. The results of this study provide no evidence that allodynia can be induced in healthy humans using a classical conditioning procedure with simultaneous timing.

Key Words. Pain; Classical Conditioning; Allodynia; Pain Threshold; Associative Learning

Introduction

Chronic pain is a societal problem: It affects 35.5% of people [1], exerts a substantial burden both financially [2] and socially [3], and according to years lived with disability, it is humanity's most burdensome health condition [4]. Despite real advances in our understanding of physiology, genetics, brain function, and comorbidities, how acute pain transitions to chronic pain remains to be elucidated. One possibility is that associative learning drives this transition [5,6]. In an acute pain episode, the simultaneous occurrence of nociceptive and non-nociceptive signaling presents a scenario that is apt to classical conditioning, a form of associative learning that occurs through the repeated presentation of a biologically evocative stimulus (unconditioned stimulus [US], e.g., a nociceptive stimulus) with an initially neutral stimulus (conditioned stimulus [CS], e.g., a tactile stimulus). After this pairing has occurred, the initially neutral stimulus comes to carry informational value about the likelihood of the biologically evocative stimulus and can evoke a response even in the absence of such a US [7]. The idea that an association would be formed between tissue damage–driven nociception and other non-nociceptive input during the acute phase of an injury seems intuitive. After healing, once nociception has diminished or ceased altogether, the informational value of the non-nociceptive input could be sufficient to drive the perseveration of pain.

Clinically, this mechanism is very well endorsed: the vast majority of clinicians believe that pain can be a classically conditioned response to non-nociceptive stimuli in the absence of nociception [8]. An extensive body of scientific literature supports the idea that nociception is not always necessary for pain (see [9] for an accessible review). Additionally, classical conditioning is known to be a dominant mechanism behind the maintenance of fear of pain [10] and increased muscular responses to nondangerous stimuli [11]—two phenomena that are commonly seen in people with chronic pain.

A recent systematic review found that, although existing pain can be *amplified* via classical conditioning mechanisms, there is insufficient evidence to support or negate the idea that pain can be *elicited* via classical conditioning [12]—in other words, that a normally nonpainful stimulus can be rendered painful (allodynia) by a classical conditioning effect. In fact, despite the idea's intuitive nature and widespread clinical endorsement, only one study reports to have elicited pain using classical conditioning [13]. However, that study could not distinguish between direct modulation of pain and a change in pain threshold secondary to changes in arousal and valence, leaving the question of classically conditioned pain unanswered.

In the current study, we aimed to use a classical conditioning paradigm to induce allodynia to a somatosensory stimulus in healthy humans. We hypothesized that pairing of one non-nociceptive somatosensory stimulus (CS+) with painful nociceptive heat (UShigh) and another (CS-) with nonpainful heat (USlow) would result in a lower pain threshold to subsequent trials of heat that were paired with the CS+ than to trials that were paired with the CS-. Our secondary aim was to investigate whether such pairing would alter subsequent perception of the CS+ relative to the CS-. We hypothesized that subjects would become more sensitive to the CS+ than to the CS- because of the difference in informational value carried by each.

Methods

Participants

We recruited healthy adult participants using flyers and word of mouth. Study information was provided electronically and verbally, and in printed form when preferred by the potential participant. Subjects were screened for exclusion criteria over the telephone or via email, and again on arrival for testing. Participants were compensated at Au$20/hour for inconvenience and travel costs. Written informed consent was obtained. All procedures conformed to the Helsinki Declaration and were approved by the institutional ethics committee.

Inclusion criteria were pain-free status, age over 18 years, and ability to consent autonomously. Screening exclusion criteria were: pain at the time of testing, use of analgesic medication on the day of testing, a current chronic pain problem, a sensation problem, diagnosed peripheral vascular disease, diabetes mellitus, a neurological problem, or a previous or current psychiatric diagnosis. Volunteers were not eligible if they had completed previous classical conditioning experiments in our laboratory. Participants were excluded if they were unable to distinguish the two tactor locations from each other (see the *Manipulation Check* section below). Midprocedure exclusion criteria were 1) thermal pain threshold greater than 50°C, 2) fewer than 75% of high-temperature thermal stimuli (UShigh) rated as painful during the acquisition phase, and 3) more than 25%

of low-temperature thermal stimuli (USlow) rated as painful during the acquisition phase (see below for details). Participants who were excluded via these criteria were not considered further. An estimated sample size was computed using G*POWER (version 3.1.9.2, Heinrich Heine Universität Düsseldorf, Germany) [14]. This calculation showed that a sample size of 34 would provide a power of 0.80 to detect a medium-sized effect when alpha was fixed at 0.05.

Stimuli

Thermal stimuli were delivered using a 900 mm^2 Peltier contact thermode controlled by Pathway software (Medoc Ltd, Ramat Yishai, Israel). The thermode was held in contact with the skin by a fabric bandage. Vibrotactile stimuli were delivered using two adapted mobile phone vibrators (tactors) that were manually controlled by software developed in house, using LabVIEW. The tactors were fixed to the skin of the back using double-sided tape (see Figure 2). Prior to the procedure, participants were allocated to receive one of the tactors as the CS+ and one site (cephalad or caudad) as the CS+ location. The tactors were set to vibrate at a clearly perceptible, nonpainful intensity for 500 ms. In case there were idiosyncratic characteristics of the stimulus delivered by each tactor, or differential effects of location, which tactor went where was counterbalanced across participants in the following sequence: 1) tactor 1, location 1; 2) tactor 2, location 1; 3) tactor 1, location 2; 4) tactor 2, location 2.

Stimulus Timing

Although forward pairing (CS presented before US) is most common in classical conditioning paradigms, we opted for a simultaneous pairing design in which the onset of the theoretical US (nociception) occurs simultaneously with the onset of the CS (non-nociceptive input). This was chosen because we deemed simultaneous timing to most closely mimic the real-life scenario of acute tissue damage, in which non-nociceptive input is functionally simultaneous with nociceptive input. Notably, this simultaneous pairing design should rule out the conventional explanation for a classical conditioning effect—that any change in responding can be attributed to a real-time modulatory effect of expectation that arises due to the presentation of the CS+. Conversely, in this simultaneous pairing design, attributing any modulatory effect of the CS+ to a difference in real-time expectation would be implausible.

Manipulation Checks

Participants' ability to distinguish the two tactor locations could critically influence the effectiveness of differential conditioning. Perceptual discrimination between the two vibrotactile stimuli was therefore tested with several randomly ordered activations of each tactor, with participants being asked to state whether the activated tactor was the one closer to the head or the one closer to the feet. The procedure only continued when the experimenter was satisfied that the participant could distinguish the two locations consistently, and participants who failed this assessment were excluded.

Considering that repeated exposure to thermal stimulation can cause habituation or sensitisation [15–18], there was a risk that the intended classical conditioning procedure could be confounded by a change in perceptual experience of the UShigh and/or USlow during the acquisition phase. This was controlled for by midprocedure exclusion criteria 2 and 3 (see the *Participants* section above), which ensured that all participants who completed the procedure had experienced at least 75% reinforcement of CS+ with painful nociception and CS− with nonpainful heat.

Outcomes and Measures

Pain Threshold

The primary outcome was pain threshold, measured using the method of limits program supplied by the Pathway software that controlled the thermal device.

Secondary outcomes were as follows:

Perceived Intensity

Perceived intensity of CS+, CS−, and US was measured using the recently developed Fifty Either Side of Threshold Numerical Rating Scale (FESTNRS). This scale was developed because we have had significant challenges with using preexisting scales to investigate pain thresholds. We required a scale that 1) allows participants to rate both nonpainful and painful experiences on the same scale, avoiding contamination between reports, 2) precisely locates the pain threshold, and 3) is straightforward to explain to participants and straightforward and intuitive for them to use. The FESTNRS defines anchors of "no sensation" at −50, "the exact point at which what you feel transitions to pain" at 0, and the "most intense pain you can imagine at +50. For the purposes of the current study, the FESTNRS offers substantial advantages over a 0–100 NRS for pain—it allows participants to report nonpainful (between −50 and 0) and painful (between 0 and +50) experiences on the same scale, with a pain threshold clearly defined at 0, and maintains the properties of a conventional pain scale for the 0–50 section of the scale. Ratings on this scale have been found to have a strong, approximately linear relationship to objective stimulus intensity ($R = 0.64$) and to electrodermal responses ($R = 0.49$) (unpublished data). In the present study, a visual version of the scale (see Figure 1) was used to reinforce the meaning and correct use of the scale ranges by marking the range of -50 to 0 as "nonpainful" and the range of 0 to +50 as "painful." Participants were allowed to

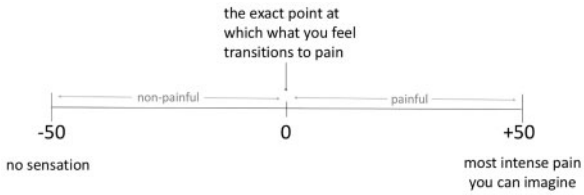

Figure 1 Visual version of the fifty either side of threshold numerical rating scale (FESTNRS) used in this study.

use any number within the range of the scale except 0. Participants were instructed to make an initial decision about whether the stimulus was painful or nonpainful and then to use the appropriate side of the scale to rate the intensity of the experience, with leftward changes in rating always indicating a decrease in intensity. They were also given examples of how the scale might be used and coached to use it proficiently (coaching materials are shown in Appendix 2). Using the FESTNRS allowed us to extract information in binary form (painful or nonpainful) without losing the sensitivity that each continuous side of the NRS provides.

Depression and Anxiety

Depression and anxiety were measured with the Depression Anxiety Stress Scale (DASS) [20]. The DASS subscales have shown high correlation with other scales measuring similar constructs and good internal consistency in several clinical groups (DASS-Depression to Beck Depression Inventory [BDI]: $r = 0.77$, Cronbach's $\alpha = 0.97$; DASS-Anxiety to Beck Anxiety Inventory (BAI): $r = 0.84$, Cronbach's $\alpha = 0.92$; DASS-Stress to BDI: $r = 0.62$; and DASS-Stress to BAI: $r = 0.64$, Cronbach's $\alpha = 0.95$ [21]). The psychometric properties are acceptable for use in research in groups with chronic pain [22], and the scale has good consistency in a general (nonpatient) adult population in South Australia (Cronbach's α for DASS-D, DASS-A, and DASS-S were 0.95, 0.88, and 0.94, respectively [23]).

Negative Affect

Negative affect at the time (i.e., state) was measured with Positive and Negative Affect Schedule (PANAS) [24], the subscales of which have shown good internal consistency (Cronbach's α for PA and NA scales were 0.89 and 0.85, respectively) and construct validity (correlations with DASS-D, Hospital Anxiety and Depression Scale (HADS)-Depression, DASS-A, HADS-Anxiety, and DASS-S, respectively, were $r = -0.48$, -0.52, -0.30, -0.31, and -0.31 for the PA subscale and 0.60, 0.44, 0.60, 0.65, and 0.67 for NA subscale) in a general adult population [25].

The Extent to Which Participants Habitually Engage in Catastrophic Thinking About Pain

The extent to which participants habitually engage in catastrophic thinking about pain was measured with the Pain Catastrophizing Scale (PCS) [26]. When tested in a student sample, the PCS has shown good internal consistency (Cronbach's $\alpha = .87$) and test-retest reliability ($r = 0.70$–0.75) and moderate correlations with measures of anxiety ($r = 0.32$) and NA ($r = 0.32$) [26]. Psychometric data were similar in a sample of community adults: Cronbach's α was 0.95, moderate scores on the PCS correlated with NA ($r = 0.31$), and PCS total scores correlated with an index of perceived pain interference ($r = 0.57$)[27].

Postexperiment questions were used to explore participants' perceptions. All questions were phrased as openly as possible, so as to avoid response biases. Some were therefore followed up with careful inquiry so that we could be certain of reaching an accurate conclusion about participants' perceptions. The first question assessed the effectiveness of blinding with, "Please explain any ideas you have about what we are testing in this study." If participants gave any indication that they might be unblinded, the question was followed up with further inquiries, so as to render this blinding check as thorough and conservative as possible. The second question assessed for any subtle differences between the qualities of vibrations at the two locations that could have contributed to learning, asking, "Did you notice any difference between the vibrations you received at the two locations, or were they identical? Please describe anything you noticed." The third question was a rudimentary assessment of the perceived similarity between the different phases of the procedure. This was included because we were aware that the perceptual experience of threshold and acquisition phases was different and that if participants conceptually distinguished between the phases, transfer of learning could be compromised. We explored this with the prompt, "Do you think that the different phases of this experiment were related, or do you think that they were independent?" The fourth question investigated the perceived timing of the thermal stimulus peaks relative to the vibratory stimuli because there was a risk that the intended simultaneous pairing, using CS onset as US reached 45°C (see the *Acquisition Phase* section below), would be perceived as backward pairing. This phrasing was, "We are interested in how you experience the timing of the *peaks* of the thermal stimuli relative to the vibratory stimuli. Did you perceive each thermal stimulus to *peak* 1) before OR, 2) at the same time as OR, or 3) after the vibratory stimulus?" The fifth question assessed contingency awareness with the phrasing, "Do you think that anything about the thermal stimulus was dependent upon the location of the vibration? If you do, what aspects of the thermal stimulus do you think were dependent?" This question was also carefully explored as necessary.

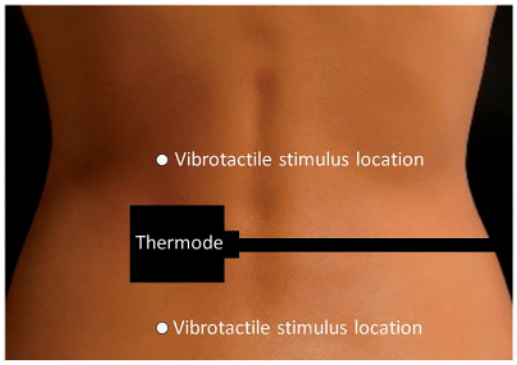

Figure 2 Layout of stimulus probes on the back.

Experimental Procedure

This study used a within-subjects differential conditioning paradigm, for which participants attended a single testing session. Participants first read through the study information sheet and had an opportunity to ask questions. They were given no precise information about the purpose of the study so as to optimize blinding. They then completed consent forms and pretest questionnaires. Testing was performed with participants lying prone on a plinth. The tactors were placed approximately 70 mm to the left of the thoracolumbar spinous processes, 70 mm apart. The thermode was positioned midway between the tactors (see Figure 2). During each test phase, the thermode was strapped to the skin using a fabric bandage. In addition to the "abort" option of verbal withdrawal from the study and the Pathway system's emergency button, participants were given a rope that allowed them to pull the thermode off the skin if the temperature became intolerable. This was an ethical necessity in view of other research groups having anecdotally reported failure of this type of system to respond to the emergency button.

The procedure began with a training period during which participants became accustomed to the thermode and could practice the task on which the threshold measurements would be based during the experiment. Participants were exposed to six trials following the "method of limits" program provided by the Pathway manufacturer. This program delivers rising heat stimulation until the participant presses a button to stop the heating, whereupon the thermode rapidly cools back to the baseline temperature. The baseline temperature of the thermode was 32°C, and its temperature rose at a rate of 1°C/second. The participant was instructed to press the button when the stimulus reached "0" (pain threshold) on the FESTNRS. The intertrial interval (ITI) was randomly jittered between 10 and 15 seconds. Scripts are provided in Appendix 1.

The experiment itself comprised five steps (see Figure 3). The order of stimuli was randomized for each phase.

Pre-Acquisition Phase

Baseline measures of the perception of the CS+ and CS− were obtained. Participants received three trials of CS+ and three trials of CS− vibrotactile stimuli in a prerandomized order and rated their experience of each stimulus on the FESTNRS. ITI was randomly jittered between 10 and 15 seconds.

The "method of limits" program was used for the identification of the pain threshold as described for the training period, except that seven stimuli were delivered so that the first could be discarded (because piloting showed the first trial to be inconsistent with subsequent trials). Each stimulus except the first was paired with a vibrotactile stimulus delivered when the thermode reached a specific temperature. The appropriate timing for delivery of the vibrotactile stimulation was based on the threshold values from the training period. If the training phase threshold values exceeded 46°C, then the vibrotactile stimuli were delivered when the thermode reached approximately 45°C. This was preferred because the simultaneous timing design required that CS and US be coincident, and the UShigh was operationalized as a painful nociceptive stimulus, based on a nociceptive activation threshold of 45°C [28]. However, if the threshold values did not reach 46°C, then the temperature at which the vibrotactile stimulus would be delivered was estimated on a case-by-case basis. The priority was to maximize the chances of a vibrotactile stimulus being delivered during each trial, without that vibrotactile stimulus being paired with too cool a thermal stimulus. As such, the temperature was usually about 1°C below the final threshold measurement determined in the training phase. The order in which vibrotactile stimuli were delivered in this phase was restricted such that pairs of CS+ and CS− were delivered sequentially, but whether CS+ or CS− was delivered first was pre-allocated at random. This order was then repeated for the threshold test 2 phase. Participants did not provide any verbal reports during this phase.

Acquisition Phase

The acquisition phase comprised 12 high-temperature stimuli (51°C, UShigh) paired with the CS+ and 12 low-temperature stimuli (45°C, USlow) paired with the CS-, delivered in a pre-randomized order, according to the Pathway software's "pulses" program. Peak temperatures were determined on the basis of pilot testing and knowledge of nociceptor activation thresholds, as discussed above (in the *Pre-Acquisition Phase* section) [28]. The thermode began at a baseline temperature of 28°C, ramped to a target temperature (in this phase, 45°C or 51°C), sustained that target temperature for one second, and then immediately cooled back to the baseline. Heating and cooling occurred at the maximum rate permitted by the hardware: 8°C/second. CS and US (high or low) needed to have approximately simultaneous onsets. The UShigh was operationalized as a

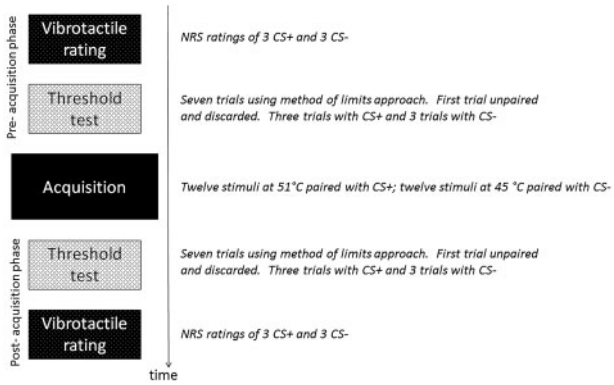

Figure 3 Experimental procedure comprising five steps in three phases: pre-acquisition, acquisition, and postacquisition.

painful nociceptive stimulus. Pain thresholds vary between individuals, but nociceptors are thought to be active at temperatures at and above 45°C [28]. Therefore, each vibrotactile stimulus was timed to coincide with the thermode reaching approximately 45°C, so that the onset of CS was approximately simultaneous with the onset of nociception. The ITI was randomized within the range of 15 to 25 seconds. Participants were asked to rate the intensity of each thermal stimulus on the FESTNRS.

Postacquisition

This phase included a second threshold test and FESTNRS ratings of vibrotactile stimuli, which were identical to the pre-acquisition phase. After testing, participants completed the postexperiment questions, they were thanked for their participation, and they left.

Statistical Analyses

Our primary question concerned the thermal pain threshold when thermal stimuli were paired with the CS+ or the CS− after the acquisition phase. The postacquisition phase was designed as an exact replica of the pre-acquisition phase, with tactile stimuli delivered in the same order and with the same timing in both phases. In order to retain this matching, errors were accounted for by eliminating the data from the matching trials (total of seven pairs of trials eliminated for this reason). In such cases, mean threshold values were calculated from the remaining matched trials. Data were visually inspected for distribution, and tests were applied to confirm that the data met the applicable test assumptions, for example, normality and sphericity for repeated measures (RM) analysis of variance (ANOVA). We analyzed the data using two RM ANOVAs. The primary analysis consisted of averaging the pain threshold results for each time point (pre-acquisition and postacquisition) and each condition and entering them into a 2 (time: threshold test 1 vs threshold test 2) x 2 (condition: CS+ vs CS−) RM ANOVA. The secondary analysis consisted of averaging the three vibrotactile stimulus FESTNRS ratings for each time point (pre-acquisition and postacquisition) and each stimulus type and conducting a 2 (time: vibrotactile pre-acquisition test vs postacquisition test) x 2 (stimulus: CS+ vs CS−) RM ANOVA.

Although we were not testing hypotheses about the relationship between gender or psychological factors and conditioning, we did explore these relationships with a view to hypothesis generation. To explore the relationship with gender, we used analysis of covariance (ANCOVA) with the same design as for the primary analysis, but with gender included as a covariate. To explore relationships with psychological factors, we used four separate regression analyses. We computed an index of the "conditioning effect" by dividing the mean threshold for CS+ test trials from the mean threshold for CS− test trials for pre-acquisition and postacquisition phases separately, and then subtracted this figure for pre-acquisition from that for postacquisition. We then used this index of conditioning as a dependent variable and used negative affect score on the PANAS, depression score on the DASS, anxiety score on the DASS, and total pain catastrophizing score on the PCS as independent (predictor) variables. We did not correct for multiple measures because these were hypothesis-generating analyses.

Results

Overview

Forty-seven participants volunteered. Four volunteers met exclusion criterion 2 (see the *Participants* section above), seven met exclusion criterion 3, one was excluded for reporting all CS as painful, and one was excluded because of equipment malfunction during testing. All excluded volunteers were replaced with new, naïve participants according to the established criteria so that full data sets were obtained for 34 remaining participants (N = 17 female participants, median

Table 1 Descriptive data for full sample (N = 34). Note that, for FESTNRS scores, positive values indicate painful experience, whereas negative scores indicate nonpainful experience.

Outcome	Measure	Mean ±SD
Negative state affect	PANAS	13.58 ± 3.23
Depression	DASS	1.53 ± 2.80
Anxiety	DASS	3.26 ± 3.53
Stress	DASS	6.44 ± 6.24
Pain catastrophizing	PCS	14.41 ± 8.39
Acquisition: intensity of high-temp stimulus	FESTNRS	23.55 ± 11.16
Acquisition: intensity of low-temp stimulus	FESTNRS	−28.42 ± 10.76

Table 2 Pain threshold results for pre-acquisition and postacquisition tests.

Pain threshold (°C)	To CS+ (mean ±SD)	To CS− (mean ±SD)
Before acquisition	44.13 ± 3.23°C	43.95 ± 3.27°C
After acquisition	46.39 ± 2.06°C	46.23 ± 2.05°C

age = 23 years, range = 18–50 years). Ratings of vibrotactile stimuli were not available for one participant, so the results for ratings of CS-only trials come from 33 participants. Questionnaire results are provided in Table 1.

Manipulation Check

All participants were able to distinguish the two tactor locations from one another prior to the conditioning procedure. Exclusions due to criteria 2 and 3 are reported above.

Primary Analysis: Pain Threshold

There was no difference between pain thresholds associated with CS+ and those associated with CS− (Table 2, no significant effect of condition $F(1,33) = 3.89$, $P = 0.057$). Pain thresholds increased from pre-acquisition to postacquisition (main effect of time, $F(1,33) = 36.45$, $P < 0.001$, $\eta^2_P = 0.525$), reflecting a habituation effect. The time x condition interaction did not reach statistical significance ($F(1,33) = 0.004$, $P = 0.949$), indicating that the acquisition procedure did not elicit a classical conditioning effect on pain threshold.

Table 3 Ratings of vibrotactile stimuli during pre-acquisition and postacquisition tests; note that the FESTNRS anchors are such that −50 denotes no sensation and 0 denotes the pain threshold, so that negative scores indicate nonpainful experience and, e.g., −34 is less intense than −33.

Rating of vibrotactile stimulus (FESTNRS)	To CS+ (mean ±SD)	To CS− (mean ±SD)
Before acquisition	−34.75 ± 8.22	−33.19 ± 8.85
After acquisition	−34.82 ± 9.56	−35.03 ± 10.31

Secondary Analysis: FESTNRS Ratings of Vibrotactile Stimulus

There was no difference between FESTNRS ratings of the CS+ and the CS− (Table 3, no significant effect of condition, $F(32,1) = 0.904$, $P = 0.349$). Ratings did not change overall from pre-acquisition to postacquisition (no main effect of time $F(1,32) = 0.328$, $P = 0.571$), but ratings of the CS− showed a greater decrease from pretest to post-test than ratings of the CS+ (significant Time x Condition interaction, $F(1,32)=4.45$, $P = 0.043$, $\eta^2_P = 0.122$) (difference between means at time 2, mean ± SD = 0.21 ± 5.03).

Exploratory Analyses

We explored the effect of gender by entering gender as a covariate in the primary ANOVA. This analysis did not show a significant time x condition interaction ($P = 0.81$), strengthening the idea that the acquisition procedure did not elicit a classical conditioning effect on pain threshold and showing that gender did not significantly influence the primary results (no interaction of time x condition x gender, $P = 0.78$). We also explored the effects of negative state affect, depression, anxiety, and pain catastrophizing using regression analyses, with the index of the "conditioning effect" (see the *Methods: Statistical Analyses* section) as the dependent variable. None of the four exploratory variables significantly predicted the strength of conditioning. The results for one-tailed tests of significance were: negative state affect $P = 0.21$; depression $P = 0.29$; anxiety $P = 0.38$; pain catastrophizing $P = 0.13$. Postexperiment questions revealed that all 34 participants were blinded to the purpose and hypothesis of the study. Four participants reported that the CS+ was consistently paired with high-temperature stimulation during the acquisition phase; 30 were unaware of this pairing. Thirty participants reported perceiving the thermal stimulus to peak after the vibrotactile stimulus; one participant reported uncertainty, saying the thermal stimulus had peaked either at the same time as the vibrotactile stimulus or before it; one participant said it had peaked before the vibrotactile stimulus, and two participants could not remember.

Discussion

We hypothesized that pairing of one non-nociceptive somatosensory stimulus (CS+) with painful heat and another (CS−) with nonpainful heat would result in a lower pain threshold to subsequent trials of heat that were paired with the CS+ than to trials paired with the CS-. Our hypothesis was not supported: we found that participants' pain thresholds were no different when, during the test phase, the threshold test was paired with the CS+ compared with when it was paired with the CS-. However, this difference between pain thresholds with the CS+ and the CS- approached significance ($P = 0.057$). Importantly, though, the critical interaction effect was absent ($P = 0.949$), and the location of the CS+ and CS- was randomly counterbalanced between participants, which excludes a difference due to skin location. Overall, our results provide no evidence that allodynia can be induced in healthy humans using classical conditioning.

Our results are in line with those of Bräscher [29], who asked a similar question using a different design. In that study, warmth stimuli (as CS+ and CS-) were paired with painful or nonpainful heat stimuli (as US). In the test phase, each conditioned stimulus was presented with the nonpainful heat stimulus. Analysis of data from the whole sample (see 12) showed that classical conditioning did not cause nonpainful thermal stimuli to become painful when they were paired with the CS+. However, this analytical approach was underpowered to detect an effect, and the trend toward a significant effect implies that replication of this design is warranted.

Our results contrast with those of Williams and Rhudy [13], who found a lower threshold when the threshold test was presented with the CS+ than when it was presented with the CS-. The contrast could be due to several differences in study design. First, the primary aim of the study by Williams and Rhudy was related to emotions: they investigated preparedness theory [30] by inducing increased arousal and negative valence to a facial expression of fear (as CS+) and then tested the effects of that increased arousal and negative valence on pain threshold. Second, during Williams and Rhudy's acquisition phase, the CSs were displayed for 10 seconds before delivery of the US. During the subsequent test phase, CSs were presented without reinforcement, and threshold measures were taken five seconds after onset of the CS. With this "delay" timing, participants presumably had enough time to identify the CS+ and would have been anticipating the US when the pain threshold test occurred. The drop in pain threshold with the CS+ may, therefore, have been mediated by the conditioned change in arousal and valence—indeed, conditioned fear is what the authors considered to have been the driver of the drop in pain threshold. In contrast, our study used neither emotional manipulation nor delay timing: we intentionally selected emotionally neutral CS and used simultaneous timing, such that the CS and US were coincident. Simultaneous timing is not commonly used in classical conditioning study designs—presumably because studies that measure anticipatory or preparatory responses to index associative learning have shown a less potent learning effect with simultaneous pairing than with forward pairing [31,32]. Of course, one would not expect an organism to develop an anticipatory behavioral response to a CS that does not temporally predict the US because such a response would serve no biological purpose [33]. However, several carefully designed studies have interrogated the assumption that simultaneous pairing is an ineffective driver of associative learning [33–35]. These studies used two-phase conditioning procedures to circumvent the limitations of anticipatory responses for detecting associations formed between simultaneously paired stimuli. Their combined results show that simultaneous pairing can, in fact, be an effective driver of conditioning and can even sometimes result in stronger associations than forward pairing [33–35]. These studies suggest that the historical exclusive measurement of anticipatory responses may have generated a false impression that simultaneous pairing produces inferior associative learning. The present study did not measure anticipatory responses, so it was reasonable for us to use simultaneous pairing. In fact, we considered simultaneous timing to be preferable because it 1) has greater ecological validity—it mimics the real-world relationship between the proposed CS (e.g., tactile input) and US (painful nociception) and 2) minimizes the likelihood of effects being driven by an indirect pathway, mediated by conditioned increases in valence or arousal. In other words, we chose to use simultaneous timing so as to target mechanisms that are not dependent on time-consuming processes. With the background of the study by Williams and Rhudy, our study's findings could be interpreted as suggesting that conditioned changes in arousal or affect—or other timing-dependent constructs like expectation—may be necessary for modulation of pain threshold (for a review, see [36]), although a previous study by our group has demonstrated classically conditioned hyperalgesia in a simultaneous timing design [37].

Our secondary hypothesis was that subjects would become more sensitive to the CS+ than to the CS- because of the difference in informational value carried by each. We found that participants reported experiencing the CS+ as more intense than the CS- after acquisition, but this difference, although statistically significant, was too small to be clinically meaningful, at 0.21 on the 100-point FESTNRS. Admittedly, participants seemed to find the task of rating the CS somewhat arbitrary. This may be because these CS-only trials were delivered in a block, rather than intermixed with other paired trials and because they did not involve a range of different CS intensities. That kind of exposure would constitute a range of experience within which to contextualize the FESTNRS. Intermixing CS-only trials with paired trials or providing a range of CS intensities could be useful

strategies to improve the sensitivity of this assessment in future designs.

One interesting finding of this study was that 30 of the 34 participants were unaware of the pairing of CS+ with painful heat. This seems surprising in light of the supposed biological relevance of painful heat, but may be a consequence of the experimental setting. Although participants had the option of terminating their involvement in the study, they did not have any choice about the order in which they received stimuli, so there was no "survival advantage" to learning about the association. Participants were also not instructed to learn during the procedure. There is considerable debate about the importance of contingency awareness for classical conditioning, and most of the evidence suggests that conditioning effects are greater when participants are aware of contingencies [38,39]. However, establishing contingency awareness as the cause of conditioned responding is challenging, given that contingency awareness might be established in parallel with conditioned responding. Elegant work by Jensen et al. has recently demonstrated classically conditioned hyperalgesia in participants to whom CS were presented subliminally, such that participants were unable to recognize the CS [40,41]. Those findings suggest that conditioning can occur without conscious awareness. However, even subliminal presentation of stimuli may lead to the establishment of expectancies. We questioned our participants about contingency awareness rather than about expectancy, so we cannot comment on whether or not they expected more pain with CS+ than with CS-. We are also unable to speculate as to whether our participants made any nonconscious associations because we did not use physiological measures in this study. Together, these considerations suggest that participants' failure to learn the association could underlie the lack of effect in the present study and that physiological measures may be a useful manipulation check to be included in future designs. Indeed, recent data suggest that a situation in which participants are aware of the contingencies may be more representative of clinical reality than a situation in which participants are unaware of the contingencies: most patients with acute back pain are able to accurately identify situations that trigger their pain [42]. This may be a particularly important consideration for studies aiming to elucidate mechanisms that could contribute to the maintenance of pain beyond the acute phase of injury.

Participants' ability to distinguish the locations of the various stimuli may also influence contingency learning. We assessed our participants' ability to distinguish the CSs from one another, but we did not assess their ability to distinguish the location of each conditioned stimulus from that of the US. A recent study by our group [37] found classically conditioned hyperalgesia and better contingency awareness after a procedure that applied stimuli, in a similar layout, to the forearm, which some studies have found to have higher spatial acuity than the back [43]. In that study, stimuli were also spaced slightly further apart, and the US was an electrical stimulus that probably provided a more spatially precise perceptual experience than the thermal stimulus used in the present study. It is possible, therefore, that spatial distinctiveness of CS and US (or of the two different CS-US combinations) facilitates contingency learning.

Another noteworthy finding is that postexperimental questioning revealed that most participants perceived the thermal stimulus to peak after the vibrotactile stimulus. This is unsurprising, considering that the vibrotactile stimuli were timed to begin when the thermal stimuli reached 45°C, which was well below the peak of the UShigh (51°C). At first glance, this seems to suggest that the present study did not succeed in achieving "simultaneous" timing—and certainly, this apparent mismatch reflects the difficulties inherent to achieving simultaneity with two different stimulus modalities. However, the present study operationalized the US as painful nociception. The rapid ramp rate of the thermode (8°C/sec) meant that the vibrotactile stimulation (duration 500 ms) would have ended as the thermode reached 49°C, which exceeds the pain threshold in most participants and certainly exceeds the nociceptor firing threshold [28]. Vibrotactile stimulation therefore began as nociception began and ended after the thermode reached a temperature that would have been painful in most participants.

The findings of this study are strengthened by effective blinding. All 34 participants were blinded to the purpose of the experiment, as confirmed at postexperiment interview. In our recent systematic review of classical conditioning of pain, blinding had not clearly been confirmed in any of the eligible studies. Indeed, one clear suggestion of that review was to reduce the risk of bias in conditioning studies by aiming for and assessing blinding of participants [12].

There are some limitations of the current study that need to be noted, and the interpretation of our results should consider two procedural factors. First, in the present study, the CS had a duration of 500 ms and was timed to begin when the threshold test stimulus reached approximately 2°C below the peak temperature attained during threshold test practice trials. This meant that the CS began and ended well before the threshold was reached in many of the test trials. It is difficult to speculate about the consequences of this timing discrepancy because so little is known about the neuronal mechanisms that might contribute to classically conditioned modulation of pain, but it is plausible that any modulating effect of the CS had been lost by the time the test stimulus approached pain threshold. It is also worth noting that this "method of limits" approach to testing pain threshold allows for operant conditioning to play a role because the test stimulus ends as soon as the participant presses the button. The safety rope that participants held to allow them to pull off the thermode in case of malfunction could also be construed to provide

a sense of control over the stimulus. However, the rope facility is similar to the opportunity for participants to verbally request cessation of the thermal stimulus or for them to press the emergency button that is routinely provided during thermal stimulation. Participants are instructed to use these only if they wish to withdraw from the study, so it is unlikely that these opportunities provided a sense of control over the intensities of individual stimuli that was sufficient to diminish the impact of the UShigh.

Second, the selection of a CS plays a fundamental role in the design of any classical conditioning study. We chose to use a somatosensory CS for its ecological validity [44] so that the paradigm would mimic the clinical scenario in which many of the signals reaching the brain would originate in the tissues close to the area from which the nociception originates. In accordance with a functional perspective on conditioning [45], we expected that somatosensory stimuli (e.g., nociception plus tactile input) would more easily be associated with each other than a somatosensory (e.g., nociception) and a nonsomatosensory (e.g., visual) stimulus might.

Third, our interpretation of the simultaneous conditioning paradigm as being capable of minimizing the effects of arousal or valence is based on timing alone. It would have been interesting to have measured CS valence in the present study, perhaps using the Self-Assessment Manikin measure [46]. Evaluative conditioning is commonly accompanied by a change in valence of the CS+, which may be attributed to the CS+ evoking the memory representation of the US (see [47] for review). However, it is also important to note that a differential change in valence of the CS would not necessarily constitute a causal relationship between that change in valence and any conditioning effect seen. Such a measure of CS valence would provide more interesting information but would not inherently support or negate the idea that simultaneous conditioning may minimize the mechanistic role of changes in arousal or valence.

It would seem sensible for further research into this question to 1) measure participants' learning—by using expectancy ratings or by measuring physiological arousal during the procedure, 2) ensure that CS are coincident with the instant at which the decision about painfulness is made during testing *or* use a different approach to assessing pain threshold, 3) consider the mechanisms by which different CS modalities might act on the painfulness of test stimuli, and 4) consider measuring the change in valence of the CS over the course of the procedure.

Importantly, this study does not provide conclusive evidence that allodynia cannot be induced using classical conditioning—only that the present carefully designed paradigm did not induce allodynia in this sample. It remains possible that allodynia might be induced under different conditions—perhaps with different stimuli [45], different anatomical sites, or different conditions of implicit threat [48].

Conclusion

The results of this study provide no evidence that allodynia can be induced in healthy humans using a classical conditioning procedure with simultaneous timing. This finding raises the possibility that previous demonstrations of classically conditioned pain modulation were reliant on conditioned changes in arousal, valence, or expectation. Alternatively, the lack of effect may be due to the lack of awareness about pairing in the majority of the participants. Finally, the current results do not support the popular idea that classical conditioning can be a driver of persistent pain in the absence of persistent nociception.

References

1 Ospina M, Harstall C. Prevalence of Chronic Pain: An Overview. Alberta, Canada: Alberta Heritage Foundation for Medical Research Edmonton, Alberta; 2002.

2 Economics A. The high price of pain: The economic impact of persistent pain in Australia. Access Economics in Collaboration with the University of Sydney Pain Management Research Institute; 2007. http://apo.org.au/node/3054.

3 Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. Eur J Pain 2006;10(4):287.

4 Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380(9859):2163–96.

5 Crombez G, Baeyens F, Eelen P. Klassieke conditionering en geconditioneerde pijn. Gedragstherapie 1994;27:97–107.

6 Moseley GL, Vlaeyen JWS. Beyond nociception: The imprecision hypothesis of chronic pain. Pain 2015;156(1):35–8.

7 Rescorla RA. Pavlovian conditioning: It's not what you think it is. Am Psychol 1988;43(3):151–60.

8 Madden VJ, Moseley GL. Do clinicians think that pain can be a classically conditioned response to a non-noxious stimulus? Man Ther 2016;22 :165–73.

9 Butler DS, Moseley GL. Explain Pain: Revised and Updated, 2nd Edition. Adelaide, Australia: Noigroup Publications; 2013.

10 Meulders A, Vansteenwegen D, Vlaeyen JWS. The acquisition of fear of movement-related pain and associative learning: A novel pain-relevant human fear conditioning paradigm. Pain 2011;152(11):2460–9.

11 Klinger R, Matter N, Kothe R, et al. Unconditioned and conditioned muscular responses in patients with chronic back pain and chronic tension-type headaches and in healthy controls. Pain 2010;150(1):66–74.

12 Madden VJ, Harvie DS, Parker R, et al. Can pain or hyperalgesia be a classically conditioned response in humans? A systematic review and meta-analysis. Pain Med 2016;17: 1094-1111.

13 Williams AE, Rhudy JL. The influence of conditioned fear on human pain thresholds: Does preparedness play a role? J Pain 2007;8(7):598–606.

14 Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39(2):175–91.

15 Breimhorst M, Hondrich M, Rebhorn C, May A, Birklein F. Sensory and sympathetic correlates of heat pain sensitization and habituation in men and women. Eur J Pain 2012;16(9):1281–92.

16 May A, Rodriguez-Raecke R, Schulte A, et al. Within-session sensitization and between-session habituation: A robust physiological response to repetitive painful heat stimulation. Eur J Pain 2012;16 (3):401–9.

17 Schmidt K, Schunke O, Forkmann K, Bingel U. Enhanced short-term sensitization of facial compared with limb heat pain. J Pain 2015;16(8):781–90.

18 Jepma M, Jones M, Wager TD. The dynamics of pain: Evidence for simultaneous site-specific habituation and site-nonspecific sensitization in thermal pain. J Pain 2014;15(7):734–46.

19 Lovibond S, Lovibond P. Manual for the Depression Anxiety Stress Scales. Sydney: Psychology Foundation; 1995.

20 Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the depression anxiety stress scales in clinical groups and a community sample. Psychol Assessment 1998;10(2):176–81.

21 Parkitny L, McAuley JH, Walton D, et al. Rasch analysis supports the use of the depression, anxiety, and stress scales to measure mood in groups but not in individuals with chronic low back pain. J Clin Epidemiol 2012;65(2):189–98.

22 Crawford J, Cayley C, Lovibond PF, Wilson PH, Hartley C. Percentile norms and accompanying interval estimates from an australian general adult population sample for self-report mood scales (BAI, BDI, CRSD, CES-D, DASS, DASS-21, STAI-X, STAI-Y, SRDS, and SRAS). Aust Psychol 2011;46(1):3–14.

23 Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. J Pers Soc Psychol 1988;54(6):1063–70.

24 Crawford JR, Henry JD. The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004;43 (3):245–65.

25 Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: Development and validation. Psychol Assessment 1995;7(4):524–32.

26 Osman A, Barrios F, Gutierrez P, et al. The pain catastrophizing scale: Further psychometric evaluation with adult samples. J Behav Med 2000;23(4):351–65.

27 Arendt-Nielsen L, Chen ACN. Lasers and other thermal stimulators for activation of skin nociceptors in humans. Neurophysiol Clin/Clin Neurophysiol 2003; 33(6):259–68.

28 Bräscher AK. Cognitive and Behavioural Context of Pain Facilitation—Nocebo Conditioning and Uncontrollability-Induced Sensitisation. Mannheim, Germany: Universität Mannheim; 2014.

29 Seligman ME. On the generality of the laws of learning. Psychol Rev 1970;77(5):406–18.

30 Pierce WD, Cheney CD. Behavior Analysis and Learning. New York: Psychology Press; 2013.

31 Savastano HI, Miller RR. Time as content in Pavlovian conditioning. Behav Process 1998;44(2):147–62.

32 Matzel LD, Held FP, Miller RR. Information and expression of simultaneous and backward associations: Implications for contiguity theory. Learn Motiv 1988;19(4):317–44.

33 Barnet RC, Arnold HM, Miller RR. Simultaneous conditioning demonstrated in second-order conditioning: Evidence for similar associative structure in forward and simultaneous conditioning. Learn Motiv 1991;22(3):253–68.

34 Rescorla RA. Simultaneous and successive associations in sensory preconditioning. J Exp Psychol 1980;6(3):207–16.

35 Wiech K, Tracey I. The influence of negative emotions on pain: Behavioral effects and neural mechanisms. NeuroImage 2009;47(3):987–94.

36 Harvie DS, Meulders A, Madden VJ, et al. When touch predicts pain: predictive tactile cues modulate perceived intensity of painful stimulation independent of expectancy. Scandinavian Journal of Pain 2016;11: 11–18.

37 Lovibond PF, Shanks DR. The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. J Exp Psychol 2002;28(1):3–26.

38 Mitchell CJ, De Houwer J, Lovibond PF. The propositional nature of human associative learning. Behav Brain Sci 2009;32(02):183–98.

39 Jensen KB, Kirsch I, Odmalm S, Kaptchuk TJ, Ingvar M. Classical conditioning of analgesic and hyperalgesic pain responses without conscious awareness. Proc Natl Acad Sci U S A 2015;112 (25):7863–7.

40 Jensen KB, Kaptchuk TJ, Kirsch I, et al. Nonconscious activation of placebo and nocebo pain responses. Proc Natl Acad Sci U S A 2012; 109(39):15959–64.

41 Parreira P, Maher CG, Latimer J, et al. Can patients identify what triggers their back pain? Secondary analysis of a case-crossover study. Pain 2016; Pain 2015;156: 1913–1919.

42 Mancini F, Bauleo A, Cole J, et al. Whole-body mapping of spatial acuity for pain and touch. Ann Neurol 2014;75(6):917–24.

43 Miguez G, Laborda MA, Miller RR. Classical conditioning and pain: Conditioned analgesia and hyperalgesia. Acta Psychologica 2014;145(1):10–20.

44 Domjan M. Pavlovian conditioning: A functional perspective. Annu Rev Psychol 2004;56(1):179–206.

45 Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry 1994;25(1):49–59.

46 De Houwer J. A conceptual and theoretical analysis of evaluative conditioning. Span J Psychol 2007;10 (2):230–41.

47 Moseley GL. Reconceptualising pain according to modern pain science. Phys Ther Rev 2007;12(3): 169–78.