Scandinavian Journal of Pain 3 (2012) 31–37

Contents lists available at SciVerse ScienceDirect

Scandinavian Journal of Pain

journal homepage: www.ScandinavianJournalPain.com

Original experimental

# Experimental pressure-pain assessments: Test–retest reliability, convergence and dimensionality

Tamara E. Lacourt*, Jan H. Houtveen, Lorenz J.P. van Doornen

*Department of Clinical and Health Psychology, Faculty of Social Sciences, Utrecht University, Utrecht, The Netherlands*

ARTICLE INFO

*Article history:*
Received 7 February 2011
Received in revised form 10 October 2011
Accepted 12 October 2011

*Keywords:*
Pain
Algometry
Pain threshold
Test–retest reliability
Principal components analysis

ABSTRACT

*Introduction:* Experimental pain studies can provide unique insight into the dimensions of pain and into individual differences in pain responsiveness by controlling different aspects of pain-eliciting stimuli and pain measures. In experimental pain studies, pain responsiveness can be assessed as pain threshold, pain tolerance or pain ratings. The test-theoretical qualities of these different measures, however, have not yet been completely documented. In the current study, several of these qualities were investigated in a pain experiment applying different algometric techniques. The objective of the study was to investigate the reliability (test–retest) and the convergent validity (correspondence) of the different methods found in the literature of measuring pressure-pain threshold, and the interrelationship between pressure-pain threshold, pressure-pain tolerance, and pressure-pain ratings.

*Methods:* Sixty-six healthy female subjects were enrolled in the study. All pressure stimuli were applied by a trained investigator, using a digital algometer with a 1 $cm^2$ rubber tip. Pressure-pain thresholds were assessed repeatedly on six different body points (i.e. left and right calf one third of total calf muscle length below the popliteal space), the lower back (5 cm left and right from the L3), and left and right forearm (thickest part of brachioradialis muscle). Next, pressure-pain tolerance was measured on the thumbnail of the non-dominant hand, followed by rating affective and sensory components (on visual analogue scales) of a stimulus at tolerance level. Last, affective and sensory ratings were obtained for two pressure intensities.

*Results:* With intraclass correlations above .75 for pain responses per body point, test–retest reliability was found to be good. However, values obtained from all first measurements were significantly higher as compared with the two succeeding ones. Convergent validity of pain thresholds across different body points was found to be high for all combinations assessed (Cronbach's alpha values >.80), but the highest for bilateral similar body parts (>.89). Finally, principal components analysis including measures of threshold, tolerance and pain ratings yielded a three-factor solution that explained 81.9% of the variance: *Moderate-level stimulus appraisal & pain tolerance; Pain threshold; Tolerance-level stimulus appraisal*.

*Conclusion and implications:* Findings of the current study were used to formulate recommendations for future algometric pain studies. Concerning pressure-pain threshold, it is recommended to exclude first measurements for *every* body point from further analyses, as these measurements were found to be consistently higher compared with the following measurements. Further, no more than two consecutive measurements (after the first measurement) are needed for a reliable mean threshold value per body point. When combining threshold values of several body points into one mean-aggregated threshold value, we suggest to combine bilateral similar points, as convergent validity values were highest for these combinations. The three-factor solution that was found with principal components analyses indicates that pressure-pain threshold, subjective ratings of moderate intensity stimuli, and subjective ratings of the maximum (tolerance) intensity are distinct aspects of pain responsiveness. It is therefore recommended to include a measure of each of these three dimensions of pain when assessing pressure pain responsiveness. Some limitations of our study are discussed.



## 1. Introduction

Experimental pain studies can provide unique insight into the dimensions of pain and into individual differences in pain responsiveness by controlling different aspects of pain-eliciting stimuli and pain measures. In experimental pain studies measures based

DOI of refers to article: 10.1016/j.sjpain.2011.11.010.

* Corresponding author at: Department of Clinical and Health Psychology, Faculty of Social Sciences, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands. Tel.: +31 302539244; fax: +31 302534718.
*E-mail address:* t.e.lacourt@uu.nl (T.E. Lacourt).

1877-8860/$ – see front matter 
doi:10.1016/j.sjpain.2011.10.003

on the stimulus intensity are frequently used, such as pain threshold (i.e. the stimulus intensity at which a persons first experiences pain) and pain tolerance (i.e. the stimulus intensity at which a person perceives the pain as unbearable). Pain responsiveness can also be assessed by pain ratings, for example pain intensity rating on a visual analogue scale. This diversity in experimental measures of pain responsiveness raises questions about the validity and reliability of these measures and their interrelationship. More insight in these issues is needed to raise experimental pain studies to a higher level.

In addition to a diversity in stimulus intensity and pain measures, different modalities of pain stimuli have also been used across studies, such as heat and cold stimuli, electrocutaneous stimuli and pressure stimuli. These stimuli do not elicit the same pain responses [1] and outcome measures between these modalities are not always related [2]. Among these modalities, pressure (i.e. algometry) has most frequently been used for comparing pain perception between pain patients and (healthy) controls [3–6]. The use of algometry in the diagnosis of rheumatic diseases indicates that pressure pain is considered a relevant experimental model for clinical pain experienced by these patients [7–9]. Other studies have demonstrated that pressure-pain perception measures are also related to clinical pain in fibromyalgia patients and in women with a wide range of pain complaints, while other stimulus modalities are not [10–12]. Because pressure pain is a modality of which the clinical relevance has been demonstrated, the focus of the current validity and reliability study will be on algometric measures.

Experimental pressure-pain studies aimed at demonstrating differences between pain patient groups and controls often show inconsistent results [13–15], which can possibly be attributed to the chosen methodology of pain assessment. Although the influence of some variations in methodology on study outcomes have already been addressed [4,12,16], there remain some fundamental issues that will be addressed in the current study. First, the pressure-pain threshold on a single body point is often assessed with repeated measures. Pain threshold is then either defined as the mean-aggregated values of (some of) these measurements [3,11], or the pressure on the last measurement [17]. Specific decisions in this procedure of data analysis, however, may have large consequences for the results. It is not clear whether using repeated measures of pain threshold (or pain tolerance for that matter) on a single point yield reliable values, and if so, how many repetitions should be performed to get a robust measure of pain threshold. Although the measurement of pressure-pain tolerance could raise the same questions, this measure is not often measured repeatedly probably because of ethical considerations. Therefore, we will restrict this issue to threshold measures only.

The second issue deals with the different methodologies found in the literature in which pressure-pain thresholds are measured on multiple body parts, often bilaterally. For example, threshold is measured on both the left and the right arm and also on the left and right calf. Data are either collapsed into a mean-aggregated threshold over the bilaterally assessed body part (e.g. the left and right arms) [2] or into a mean-aggregated overall threshold including all body parts (e.g. both arms and legs) [4]. Although it has been shown that there is no statistical difference in pain threshold between dominant and non dominant site per body part [3], it is unknown what the correspondence is between pressure-pain thresholds on different body parts and whether this procedure masks area-specific effects.

The third and final issue addressed in the current study focuses on the difference between using pressure-pain thresholds and tolerance levels versus using pain ratings as the main outcome. In a typical experimental pressure-pain threshold or tolerance study, participants undergo several stimulus intensities. Next, the intensity of stimulus threshold or tolerance (in for example kilopascal) is used as outcome and compared between individuals. In a typical pain-rating study, self-reported intensity or unpleasantness of pain (using a VAS or Likert scale) of a predefined stimulus intensity is compared between individuals. However, tests for differences in pressure-pain threshold and tolerance may yield different results than tests for differences in ratings of stimuli. It has already been demonstrated that pressure-pain thresholds based on *ratings of stimulus intensity* yielded different results than pressure-pain thresholds based on *the first indication that the stimulus became painful* [4]. There is also evidence that intensity and unpleasantness ratings do relate to clinical pain measures differently than pain threshold and tolerance levels [10]. These measures may very well reflect distinct aspects of pain.

The objective of the current study was to investigate the reliability (test–retest) and the convergent validity (correspondence) of the different methods found in the literature of measuring pressure-pain threshold, and the interrelationship between pressure-pain threshold, pain tolerance, and pain ratings. Specifically, the following questions will be answered: (a) what is the reliability of repeated measurements of pressure-pain threshold on a single body point, (b) what is the correspondence between pressure-pain thresholds across different body parts, and (c) to what extent are pressure-pain threshold, pressure-pain tolerance, and pain ratings aspects of a common dimension or of separable dimensions. Based on the results of the current study, methodological recommendations for future experimental algometric pain studies will be formulated.

## 2. Methods

### 2.1. Subjects and ethical considerations

Eligible female undergraduates were selected based on their answers on an online survey on health and health behaviour. Only females were included in the sample to increase homogeneity. The online survey was advertised on posters and flyers which were distributed throughout the university campus. A total of 370 female students filled in the online survey, of which 263 gave permission to be invited for the current study. One hundred-eleven respondents were excluded based on having a medical condition (including having the flu or a cold), using prescribed medication (including frequent use of asthma medication, not including contraceptives), or both. From the 152 respondents who were invited to participate, 66 showed up. All subjects participated in exchange for remuneration or course credit. The instituional ethical review board considered the study acceptable.

### 2.2. Instruments

All pressure stimuli were applied by a trained investigator. A digital algometer (FDX 50; Wagner Instruments) was used with a 1 cm$^2$ rubber tip that was placed on the skin or finger nail. Pressure was recorded in kilopascal (kPa), with the algometer recording graduations of 1.96 kPa.

For pressure-pain tolerance, an additional holding device was designed with which the pressure on the thumb nail could be applied in a more controlled manner. The algometer was placed in a wooden casket, which could be moved up and down inside a wooden column. Participants inserted their thumb into the column underneath the algometer, after which the algometer could be pressed down manually by the investigator by moving a lever down. Safety blocks inside the column made sure the casket could not be moved down entirely, thus ensuring enough space for the thumb. This device was also used in applying discrete stimuli for assessment of pressure-pain ratings.

**Table 1**
Overview of pressure-pain measures and methods.

| Pressure-pain measures | Procedure |
|---|---|
| Pressure-pain threshold (PPth) | Three measurements on left and right calf, lower back, and forearm |
| Pressure-pain tolerance (PPtol) | Maximal tolerance of thumb pressure |
| Pressure-pain ratings | Unpleasantness and painfulness of thumb pressure maximal (tolerance), 294, and 490 kPa |

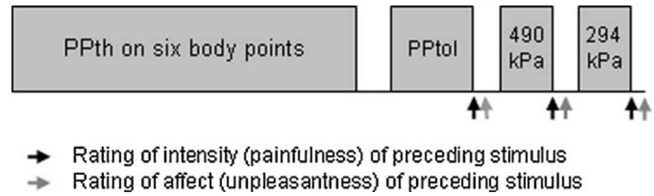


**Fig. 1.** Time line of assessment of pressure-pain measures.

### 2.3. Pressure-pain measures

An overview of the pressure-pain measures assessed in the current study is given in Table 1.

#### 2.3.1. Pressure-pain threshold

In the current study, pressure-pain threshold was defined as the pressure at which the participant first indicated the pressure to become unpleasant. The expression "unpleasant" was chosen to avoid possible anxiety effects of the expression "painful". During pressure-pain threshold (PPth) measurement, the participant was in a prone position on a massage table with the head facing down in a face rest. Pressure was applied using the algometer and gradually increased with 98 kPa per second until the participant indicated that the pressure became unpleasant by saying "stop", after which the algometer was immediately removed from the skin and maximum pressure was copied from the algometer screen. To enable the investigator to increase the pressure at a constant rate, a digital metronome was used to indicate 1 s time interval (with soft ticking sounds). The use of the metronome was explained to the participants as a tool to help the investigator with the measurements, but it was not explained how this was used.

PPth was measured three times on six body points in a fixed order without breaks: left and right calf (one third of total calf muscle length below the popliteal space), the lower back (5 cm left and right from the L3), and left and right forearm (thickest part of brachioradialis muscle) respectively. On each body point, the three measurements were taken successively before moving to the next point. Time intervals between measurements were 30–40 s. Pressure did not exceed 1471 kPa at the legs or the lower back or 980.5 kPa at the arms. When PPth was not indicated before reaching these pressures, 1471 or 980.5 kPa was noted down as the maximum pressure.

#### 2.3.2. Pressure-pain tolerance

Pressure-pain tolerance was defined as the maximum pressure at which the participant indicated the pressure to become too painful. Pressure-pain tolerance (PPtol) was assessed on the thumb of the non-dominant hand with the subject sitting in a chair using the holding device described above. Pressure was increased with 98 kPa per second until the subject indicated the pressure to become too painful. After each trial, the subject was asked whether or not that point was actually reached. If not, the subject was asked to try again. Inter-stimulus times were 30–40 s. Maximum pressure was recorded for every trial. Pressure did not exceed 1471 kPa. In case PPtol was not indicated at 1471 kPa, this pressure was noted down as PPtol.

#### 2.3.3. Pressure-pain ratings

For several stimulus intensities, two pressure-pain sensations were rated. Unpleasantness (affective aspect) of the pressure stimulus was rated on a 100 mm visual analogue scale ranging from "not unpleasant at all" to "the most unpleasant feeling ever", while painfulness (sensory aspect) was rated on a 100 mm scale ranging from "not painful at all" to "the most painful feeling ever". First, the participant was asked to rate the sensation of the pressure pain tolerance stimulus on the thumb of the non-dominant hand for the last trial of the tolerance measurement (i.e. the trial on which the participant indicated that the pressure-pain tolerance level was reached). Second, discrete stimuli of 294 kPa and 490 kPa were applied to the thumb of the dominant hand and the participant was instructed to rate the sensation of these stimuli directly after the pressure stimulus was applied. Pressure was again increased with 98 kPa per second until the desired level was reached. On reaching either 294 or 490 kPa, pressure was maintained for two seconds after which the algometer was removed from the nail.

### 2.4. Procedure

Pressure-pain measures were taken between 9 am and 5 pm in a laboratory setting. Subjects were scheduled according to their own preference and were asked to refrain from caffeine intake in the 2 h before participating. See Fig. 1 for a time line of the procedure. After obtaining informed consent, pressure pain perception was assessed in the following order: PPth, PPtol, subjective ratings of a stimulus at tolerance intensity and of discrete pressure stimuli, with short breaks (i.e. 1–3 min) between measures (see Fig. 1). Verbal instructions were given before every measurement, followed by practice trials. The protocol was in accordance with the principles of the Declaration of Helsinki (October, 2008) and approved by the ethical committee of the Faculty of Social Sciences.

At the end of the session, subjects were asked to fill out a questionnaire booklet, containing questions regarding current pain complaints, current medical condition (which started after the internet survey) and current medication use, as well as control questions regarding caffeine intake in the 2 h preceding the study, sleep during the last two nights and amount of intensive physical exercise on the day preceding the study or on the day of the study itself.

### 2.5. Statistical analysis

All measures were tested for normality and presence of outliers and, when needed, appropriate transformations (i.e. logarithmic) were applied. For analyzing test–retest reliability of the repeated measures of PPth per body point, intraclass correlation coefficients (ICC) in a two way mixed model with type 'absolute agreement' were calculated. The correspondence between PPth measures across different body parts was assessed with reliability analysis using Cronbach's alpha. To investigate to what extent pain tolerance, pain thresholds, and pain ratings entail different aspects of pain responsiveness, the factor structure of the pressure-pain measures (i.e. PPth, PPtol and subjective ratings of tolerance intensity and of 294 and 490 kPa stimuli) was investigated with principal components analysis. The number of factors was determined by examining the eigenvalues of the factors, where factors with eigenvalues $\geq 1$ were retained. Since factors were expected to correlate with each other, oblique (oblimin) rotation was performed before interpreting the factor solution. All analyses were performed with SPSS 16.0 for Windows [18].

**Table 2**
Mean pressure (kPa) for each pressure pain threshold measurement (1st, 2nd and 3rd) per body point and for pressure pain tolerance.

| Body point | Measurement | *N* | Mean (SD, SE) | Range |
|---|---|---|---|---|
| Pressure pain threshold | | | | |
| Left calf | 1st | 65 | 463.54 (156.29, 19.39) | 219.67–970.86 |
| | 2nd | 66 | 414.84 (151.40, 18.64) | 176.52–819.84 |
| | 3rd | 66 | 392.65 (158.16, 19.47) | 141.22–908.10 |
| Right calf | 1st | 66 | 476.34 (179.63, 22.11) | 217.71–1149.34 |
| | 2nd | 66 | 417.26 (151.63, 18.67) | 180.44–1017.93 |
| | 3rd | 66 | 413.24 (169.44, 20.86) | 176.52–1176.80 |
| Lower back left | 1st | 66 | 474.94 (195.93, 24.12) | 125.53–1100.31 |
| | 2nd | 66 | 449.56 (193.82, 23.86) | 123.56–1166.99 |
| | 3rd | 66 | 453.75 (202.84, 24.97) | 117.68–1212.10 |
| Lower back right | 1st | 66 | 471.88 (185.40, 22.82) | 143.17–906.14 |
| | 2nd | 66 | 454.31 (190.86, 23.49) | 166.71–963.01 |
| | 3rd | 66 | 457.23 (193.96, 23.88) | 141.22–925.75 |
| Left arm | 1st | 65 | 313.64 (117.66, 14.59) | 98.07–660.97 |
| | 2nd | 66 | 285.34 (118.88, 14.63) | 103.95–662.93 |
| | 3rd | 66 | 291.43 (133.74, 16.46) | 109.84–747.27 |
| Right arm | 1st | 66 | 307.14 (127.99, 15.75) | 121.60–862.01 |
| | 2nd | 66 | 281.92 (129.06, 15.89) | 129.45–947.32 |
| | 3rd | 66 | 286.53 (137.30, 16.90) | 94.14–819.84 |
| Pressure pain tolerance | | | | |
| Thumb nail | Last | 66 | 967.77 (259.05, 31.89) | 355.00–1465.11 |

## 3. Results

### 3.1. Subjects

A sample of 66 female undergraduates participated in the study, with a mean age of 21.53 (sd: 3.47) and mean body mass index of 21.99 (sd: 2.52). Twelve subjects (18.2% of the sample) indicated to have a cold or to have mild flu-like symptoms on the day of study participation, although none of them felt too unwell to participate. Twenty-one subjects (31.8%) reported pain complaints in one or several body parts, of which 11 indicated pain in neck, shoulders, or upper back, and 4 reported pain in the middle or lower back. Finally, some participants ($n = 9$; 13.6%) used medication on the day before participating or on the day of participation: over-the-counter pain inhibitors ($n = 4$); anti-histamine ($n = 2$); antibiotics ($n = 1$); cholesterol synthesis inhibitors ($n = 1$); and thyroid gland hormones ($n = 1$). No differences between these subgroups and the remaining sample were found on the pain measures, thus none of the subgroups were excluded.

### 3.2. Pressure-pain measures

Mean pressures for the threshold and tolerance measurements are shown in Table 2. Skewness and kurtosis data were indicative of non-normality for most threshold measures, which was confirmed by tests of normality (Kolmogorov–Smirnov was non-significant for only 4 out of 18 measurements). Log transformation resulted in normality for all measures and less outliers. The log-transformed variables of threshold measures were used for all further analyses.

On average, 2 trials were needed before participants indicated their actual PPtol level was reached ($M = 1.8$, range 1–5 trials). When more than one trial was recorded, maximum pressure on the last trial was used as PPtol for that individual. Tests of normality did not yield any deviance from normal distributions and the untransformed variables were used for all further analyses.

**Table 4**
Intraclass correlation coefficients (ICC) for repeated pressure-pain threshold measures (log transformed) per body point.

| | Repeated measurement combinations | |
|---|---|---|
| | 1st, 2nd and 3rd | 2nd and 3rd |
| Left calf | .794 | .848 |
| Right calf | .758 | .808 |
| Lower back left | .896 | .926 |
| Lower back right | .940 | .952 |
| Left arm | .862 | .933 |
| Right arm | .904 | .934 |

Mean sensory and affective ratings are given in Table 3. Skewness and kurtosis data showed a normal distribution of the sensory and affective ratings of the 490 kPa stimulus and the stimulus at tolerance level. This was confirmed by Kolmogorov–Smirnov tests (all $p$-values > 05). Although both affective and sensory ratings of the 294 kPa stimulus showed non-normality, transformations did not result in better distributions, therefore the untransformed variables were used for all further analyses.

### 3.3. Test–retest reliability of repeated measures of pressure-pain threshold

Intraclass correlation coefficients (ICC) for repeated measures per body point are shown in Table 4 (first row). These values were all above .75, indicating good test–retest reliability. However, Fig. 2 shows that first measurements on every body point yield higher thresholds compared with second or third measurements, which was confirmed in a repeated measures multivariate analysis of variance where a significant main effect for the three

**Table 3**
Mean ratings of painfulness (sensory) and unpleasantness (affective) of pressure stimuli on 0–100 mm scales.

| Stimulus intensity | Sensory rating | | | Affective rating | | |
|---|---|---|---|---|---|---|
| | *N* | Mean (SD) | Range | *N* | Mean (SD) | Range |
| Tolerance-level | 66 | 63.0 (17.1) | 18.0–98.0 | 66 | 66.4 (20.3) | 12.5–100 |
| 294 kPa | 66 | 22.4 (23.2) | 0.0–92.0 | 66 | 24.6 (89.0) | 0.0–89.0 |
| 490 kPa | 64[a] | 34.4 (24.3) | 0.0–100 | 64 | 37.9 (27.5) | 0.0–100 |

[a] For two participants pain tolerance level was below 490 kPa; these participants were excluded from rating the 490 kPa pressure stimulus.

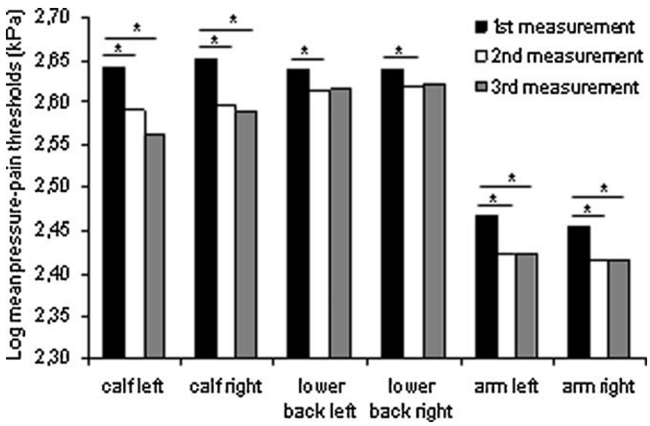


**Fig. 2.** Log mean pressure pain thresholds per measurement and body point. * Significant difference between measurements as analyzed with paired samples *t*-tests ($p < .017$).

**Table 5**
Cronbach's alpha statistics for mean pressure pain threshold (mean is calculated over log transformed data) measures across combinations of body points (bilateral, unilateral, or all).

| | Cronbach's alpha |
|---|---|
| Bilateral | |
| Leg | .890 |
| Lower back | .944 |
| Arm | .948 |
| Unilateral | |
| Left | .814 |
| Right | .872 |
| All | .931 |

measurements was found ($F(2,62) = 44.27$, $p < .0001$). Post hoc paired samples *t*-tests with Bonferoni corrected *p*-values ($p > .017$ indicated significance) indeed showed significant differences in pressure-pain thresholds between 1st and 2nd and between 1st and 3rd measurements on most body points, and no difference between 2nd and 3rd measurements on all body points. Thus, ICCs were also computed for second and third measurements only, resulting in higher ICCs for all body points (Table 2, second row). For further analyses, a mean pressure-pain threshold was calculated per body point, using the log transformed values from the 2nd and 3rd measurements only.

### 3.4. Internal consistency of pressure-pain threshold measures across body points

Internal consistency of the mean pressure-pain thresholds across body points was assessed with Cronbach's alpha. Alpha values were computed across: (a) bilateral body points (e.g. body points on the left and right arm), (b) unilateral body points (i.e. all left body points or all right body points), and (c) all body points. See Table 5 for all alpha values. These values were high (above .80) for all comparisons, but somewhat lower for unilateral combinations than for bilateral combinations or when combining all body points. Thus, results indicate that consistency is higher across bilateral same body parts as compared with unilateral combination. For further analyses, mean PPth's over bilateral body points were calculated.

**Table 7**
Pattern matrix of the principal component analysis with oblimin rotation. Only loadings >.72 are depicted.

| | Moderate level stimulus appraisal & pain tolerance | Pain threshold | Tolerance-level stimulus appraisal |
|---|---|---|---|
| PPth leg | | .913 | |
| PPth back | | .847 | |
| PPth arm | | .890 | |
| PPtol | −.818 | | |
| Sensory rating 294 kPa | .903 | | |
| Affective rating 294 kPa | .871 | | |
| Sensory rating 490 kPa | .917 | | |
| Affective rating 490 kPa | .887 | | |
| Sensory rating PPtol | | | .894 |
| Affective rating PPtol | | | .819 |

### 3.5. Interrelationship of pressure-pain measures

To investigate to what extent PPth, PPtol, and subjective ratings of tolerance level, 294, and 490 kPa are distinct aspects of pain responsiveness, principal components analysis (PCA) was performed including all pain measures. First, suitability of the data for PCA was assessed. All pain measures were checked for univariate outliers by inspecting standardized scores (*z* scores). *Z* scores above 3.0 were considered outliers. No outliers were detected in the variables. Table 6 shows correlations between all pain variables. Several correlations were .30 or above, indicating that inter-correlations were sufficiently high to be suitable for PCA. This was confirmed by Bartlett's test of sphericity, which was significant ($p < .05$) and the Kaiser–Meyer–Olkin measure, which was .74 (.60 is considered as the minimum value for a good factor analysis [19]).

PCA revealed the presence of three factors with eigenvalues exceeding 1, explaining 50.2, 20.6, and 11.1% of the variance respectively. The scree plot also pointed to a three-factor solution. All three factors showed a number of strong loadings and all variables loading substantially on only one component. Table 7 shows the pattern and structure matrix for the solution. The first factor contains PPtol and the sensory and affective ratings of the discrete stimuli, named "Moderate-level stimulus appraisal & pain tolerance" for future reference. The second factor holds all three PPth measures and will be referred to as "Pain threshold", and as the third factor comprises the ratings of the stimulus at

**Table 6**
Pearson correlations between pressure-pain measures.

| | Mean PPth leg | Mean PPth back | Mean PPth arm | PPtol | Sensory rating PPtol | Affective rating PPtol | Sensory rating 294 kPa | Affective rating 294 kPa | Sensory rating 490 kPa |
|---|---|---|---|---|---|---|---|---|---|
| Mean PPth lower back | .654 | | | | | | | | |
| Mean PPth arm | .712 | .731 | | | | | | | |
| PPtol | .396 | .359 | .442 | | | | | | |
| Sensory rating PPtol | .088 | .069 | .036 | −.084 | | | | | |
| Affective rating PPtol | −.111 | −.146 | −.207 | −.118 | .616 | | | | |
| Sensory rating 294 kPa | −.239 | −.375 | −.382 | −.654 | .456 | .438 | | | |
| Affective rating 294 kPa | −.227 | −.399 | −.439 | −.616 | .361 | .453 | .935 | | |
| Sensory rating 490 kPa | −.139 | −.285 | −.241 | −.596 | .504 | .395 | .847 | .785 | |
| Affective rating 490 kPa | −.220 | −.345 | −.356 | −.552 | .337 | .467 | .853 | .902 | .865 |

tolerance level this factor will be referred to as “Tolerance-level stimulus appraisal”. Moderate correlations were found between several factors, indicating that the oblique rotation was suitable [19]: $R^{\text{Moderate-level stimulus appraisal \& pain tolerance, Pain threshold}} = -.360$, $R^{\text{Moderate-level stimulus appraisal \& pain tolerance, Tolerance-level stimulus appraisal}} = .332$ and $R^{\text{Pain threshold, Tolerance-level stimulus appraisal}} = .036$.

## 4. Discussion

The current study was designed to examine reliability, validity and interrelationships of several measures that can be obtained in experimental pressure-pain studies. Pressure-pain threshold was measured three times on six body points. Results of the current study demonstrated good test–retest reliability for the three repeated measurements, as indicated by inter-measurement correlations of .70 or higher. However, it was also found that every first trial on a body point yielded significantly higher thresholds as compared with the second and third trials, while no difference was found between thresholds for the second and third, indicating the necessity of a practice trial on every point that is measured. Thus, practicing threshold measurements on only one body point, as was done in the current study, is not enough. Based on these findings, it is recommended to exclude first measurements for *every* body point from further analyses. Also, since no differences were found between second and third measurements, only two consecutive measurements (in addition to the practice trial) are needed.

As threshold measures on different body parts are often combined into one mean-aggregated individual PPth, internal consistency of three regularly used combinations of separate PPths (i.e. bilateral, unilateral, all) was also investigated in the current study. Although all three combinations showed good internal consistency (i.e. Cronbach’s alpha values >.80), consistency was highest for bilateral combinations (combining PPth’s of one body part bilaterally assessed, e.g. left and right arm). Thus, the correspondence between threshold-values of identical left versus right body parts is larger than the correspondence between different body parts. Although it can be concluded that mean-aggregations of all combinations of threshold are adequate, aggregation across different body parts will lead to less reliable measures as compared with bilateral same body parts. Thus, we advocate to aggregate bilateral combinations of pain threshold values.

The factor solution as assessed in the current study with principal components analysis shows a three-factor structure: *Moderate-level stimulus appraisal & pain tolerance*, *Pain threshold*, and *Tolerance-level stimulus appraisal*, with moderate correlations between factors. Based on this factor solution it is concluded that pressure-pain responsiveness indeed consists of different dimensions. Each dimension (i.e. factor) adds to the underlying construct. Studies that do not include all of these three distinct dimensions of pain may yield an incomplete picture of pain responsiveness.

In the three-factor solution, subjective ratings of moderate pain stimuli (i.e. 294 and 490 kPa) describe a different aspect of pain experience than threshold measures, and surprisingly the subjective ratings tap the same dimension as (maximum) pain tolerance (in kPa), whereas the subjective judgments of this maximal load forms a separate factor. These results may indicate that PPtol is a more subjective measure of pain responsiveness (sharing the same subjective judgment aspect with the judgment of lower intensity stimuli), while PPth measures may be considered as more objective measures of pain responsiveness. The subjective nature of PPtol is also demonstrated by the finding that the affective and sensory ratings are highly correlated at lower intensities (around .90) and at the tolerance intensity only a .60, implying that while at lower intensities sensory and affective aspects are not distinguished from each other, they are well distinguished at tolerance intensity. The high correspondence found between pain tolerance and subjective pain reports to moderate pain stimuli also implies that the tolerance for a hardly bearable painful stimulus may also be assessed by subjective reports of pain to much lighter pressure intensities, thereby avoiding discomfort in the subjects.

The strength of the relation found between affective and sensory ratings might be interpreted as that people can hardly discriminate subjectively between these two aspects of pain. There is, however, a solid physiological basis to consider affective and sensory ratings as different aspects of pain. The correspondence found in the current study may, on the other hand, also be due to the fact that a common instrument was used, i.e. the visual analogue scale on which the subjects could rate the intensity and unpleasantness of the stimuli. This explanation is supported by the findings of other studies using numerical descriptor scales to assess unpleasantness and intensity of stimuli. These studies did find differences between unpleasantness and intensity ratings [1,5], indicating a role for scale properties.

The current study can contribute to understanding some inconsistencies among the results of previous pain studies. For example, the clear distinction we have found between first and subsequent measurements of pressure-pain thresholds could explain the differences between previous studies that focused on the relations between pain thresholds and clinical pain in fibromyalgia samples. Pressure-pain threshold was found to be related to pain experienced in the present *past* but not to *present* pain in some studies [10,20], while the opposite pattern was found in another study (i.e. a relation was found with *present* pain, but not with pain in the present *past* [11]). In the first two studies mentioned, however, threshold was measured with *one* trial, while in the last study threshold was averaged over three trials. The results of the current study suggest that threshold measures averaged over multiple trials are more accurate reflections of true pain responsiveness (i.e. sensory processes), and therefore they should more reliably correspond with present pain. It may be speculated that the results obtained from a single trial threshold are confounded by psychological mechanisms. This is supported by the finding that a one-trial threshold was related to retrospectively reported pain in the past. It is known that retrospective report of complaints (such as in report of pain in present past) yields higher complaint scores compared with momentary report [21]. This discrepancy between retrospective and momentary report is probably (just as the first threshold) also related to psychological factors [22]. Thus, the relation between single-trial thresholds and pain in present past (as has been found by Geisser et al. and Giesecke et al. [10,20]) could possibly reflect some psychological factors, while the relation found between three-trial thresholds and present pain (as has been found by Lautenbacher et al. [11]) reflects true pain responsiveness (i.e. sensory processes).

Several limitations of our study have to be mentioned. First, only females within a specific range of age participated in this study. This limits the generalizability of the results with males and older subjects. A recent review shows that females have lower pressure-pain threshold and tolerance levels compared with males [23]. Also, there is some evidence that emotional and cognitive processes affect pain measures more in males than in females [24–26]. Thus, it is possible that somewhat different results will be found when the current study is replicated with a male sample. Specifically, we would expect that in males more practice trials are needed to get an accurate PPth measure by lowering anxiety levels. Concerning our recommendation to aggregate measures across bilateral body parts, we would expect this to be valid for males as well, as gender differences in pain measures are found on several different body parts [23]. Another limitation is that experimental pressure pain is often used in the assessment of (chronic) pain cohorts and it is not clear whether the results from the current study can be extended to these groups. However, the sample in our study did entail several participants with pain complaints and preliminary results showed no difference in pain measures between these participants and pain free participants, indicating that the results can possibly also be extended to pain groups. Third, several findings in our study may also be ascribed to methodological issues. The finding that ratings of unpleasantness were not distinguishable from ratings of painfulness can be explained by the similarity of the scales used. As mentioned above, several studies using different scales did find differences in outcome measures. Further, it is possible that the factor solution found in our study represents underlying methodological differences and similarities between measurements: both the PPtol and the discrete stimuli were applied on the thumb nail with the algometer placed in a casket, in sight of the subject and the subject sitting on a chair, while for measurement of PPth, the subject was lying face down on a massage table and could not see the hand-held algometer. Future studies in which the factor analysis will be expanded with subjective ratings of stimuli delivered in the same manner as PPth was assessed, would give further insight in this issue. Finally, a relatively fast incremental rate of the pressure was used in the current study (i.e. 98 kPa/s), possibly affecting the accuracy of the tests because of short response times for the subjects in cases of low pain thresholds. However, the test–retest reliability was high for all body points tested while the pressure pain thresholds differed, indicating that this was not the case.

Summarizing, the current study aimed to investigate reliability and validity of several pressure-pain methods. Despite some limitations, some recommendations for use of pressure-pain methods are formulated: first, PPth, subjective ratings of moderate intensity stimuli, and subjective ratings of the maximum intensity are distinct aspects of pain responsiveness. It is therefore recommended to include a measure of each of these three dimensions of pain when assessing pressure pain responsiveness. Second, in the assessment of PPth, we recommend one practice trial per body point and two consecutive trials for computing a mean PPth per body point. Further, when it is desirable to collapse PPth on several body points into one mean PPth, we suggest to average over bilateral body points only.

## References


[1] Rainville P, Feine JS, Bushnell MC, Duncan GH. A psychophysical comparison of sensory and affective responses to four modalities of experimental pain. Somatosens Mot Res 1992;9:265–77.

[2] Bhalang K, Sigurdsson A, Slade GD, Maixner W. Associations among four modalities of experimental pain in women. J Pain 2005;6:604–11.

[3] Maquet D, Croisier JL, Demoulin C, Crielaard JM. Pressure pain thresholds of tender point sites in patients with fibromyalgia and in healthy controls. Eur J Pain 2004;8:111–7.

[4] Chang L, Mayer EA, Johnson T, Fitzgerald LZ, Naliboff B. Differences in somatic perception in female patients with irritable bowel syndrome with and without fibromyalgia. Pain 2000;84:297–307.

[5] Petzke F, Harris RE, Williams DA, Clauw DJ, Gracely RH. Differences in unpleasantness induced by experimental pressure pain between patients with fibromyalgia and healthy controls. Eur J Pain 2005;9:325–35.

[6] Giesecke T, Gracely RH, Grant MAB, Nachemson A, Petzke F, Williams DA, Clauw DJ. Evidence of augmented central pain processing in idiopathic chronic low back pain. Arthritis Rheum 2004;50:613–23.

[7] Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS, Medsger Jr TA, Mitchell DM, Neustadt DH, Pinals RS, Schaller JG, Sharp JT, Wilder RL, Hunder GG. The American rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

[8] Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, Tugwell P, Campbell SM, Abeles M, Clark P, Fam AG, Farber SJ, Fiechtner JJ, Franklin CM, Gatter RA, Hamaty D, Lessard J, Lichtbroun AS, Masi AT. The American College of Rheumatology 1990. Criteria for the classification of fibromyalgia. Report of the multicenter criteria committee. Arthritis Rheum 1990;33:160–72.

[9] Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, Russell AS, Russell IJ, Winfield JB, Yunus MB. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 2010;62:600–10.

[10] Geisser ME, Gracely RH, Giesecke T, Petzke FW, Williams DA, Clauw DJ. The association between experimental and clinical pain measures among persons with fibromyalgia and chronic fatigue syndrome. Eur J Pain 2007;11:202–7.

[11] Lautenbacher S, Rollman GB, McCain GA. Multi-method assessment of experimental and clinical pain in patients with fibromyalgia. Pain 1994;59:45–53.

[12] Petzke F, Gracely RH, Park KM, Ambrose K, Clauw DJ. What do tender points measure? Influence of distress on 4 measures of tenderness. J Rheumatol 2003;30:567–74.

[13] O'Neill S, Manniche C, Graven-Nielsen T, Arendt-Nielsen L. Generalized deep-tissue hyperalgesia in patients with chronic low-back pain. Eur J Pain 2007;11:415–20.

[14] Schmaling KB, Hamilos DL, DiClementi JD, Jones JF. Pain perception in chronic fatigue syndrome. J Chron Fatigue Synd 1998;4:13–22.

[15] Meeus M, Roussel NA, Truijen S, Nijs J. Reduced pressure pain thresholds in response to exercise in chronic fatigue syndrome but not in chronic low back pain: an experimental study. J Rehabil Med 2010;42:884–90.

[16] Petzke F, Clauw DJ, Ambrose K, Khine A, Gracely RH. Increased pain sensitivity in fibromyalgia: effects of stimulus type and mode of presentation. Pain 2003;105:403–13.

[17] Sayed-Noor AS, Englund E, Wretenberg P, Sjödén GO. Pressure-pain threshold algometric measurement in patients with greater trochanteric pain after total hip arthroplasty. Clin J Pain 2008;24:232–6.

[18] SPSS 16.0 for Windows, Rel. 16.0.2. Chicago: SPSS Inc.; 2008.

[19] Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Boston: Pearson Education, Inc.; 2007.

[20] Giesecke T, Gracely RH, Williams DA, Geisser ME, Petzke FW, Clauw DJ. The relationship between depression: clinical pain, and experimental pain in a chronic pain cohort. Arthritis Rheum 2005;52:1577–84.

[21] Houtveen JH, Oei NYL. Recall bias in reporting medically unexplained symptoms comes from semantic memory. J Psychosom Res 2007;62:277–82.

[22] Sohl SJ, Friedberg F. Memory for fatigue in chronic fatigue syndrome: relationships to fatigue variability, catastrophizing, and negative affect. Behav Med 2008;34:29–35.

[23] Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley III JL. Sex, gender and pain: a review of recent clinical and experimental findings. J Pain 2009;10:447–85.

[24] Fillingim RB, Hastie BA, Ness TJ, Glover TL, Campbell CM, Staud R. Sex-related psychological predictors of baseline pain perception and analgesic responses to pentazocine. Biol Psychol 2005;69:97–112.

[25] Jones A, Zachariae R, Arendt-Nielsen L. Dispositional anxiety and the experience of pain: gender-specific effects. Eur J Pain 2003;7:387–95.

[26] Jones A, Zachariae R. Investigation of the interactive effects of gender and psychological factors on pain response. Br J Health Psychol 2004;9:405–18.