

# Pain rating schema: three distinct subgroups of individuals emerge when rating mild, moderate, and severe pain

Laura A Frey-Law<sup>1</sup>

Jennifer E Lee<sup>2,3</sup>

Alex M Wittry<sup>4</sup>

Myles Melyon<sup>5</sup>

<sup>1</sup>Department of Physical Therapy and Rehabilitation Science, The University of Iowa, Iowa City, IA, USA; <sup>2</sup>Mount Mercy University, Department of Psychology, Cedar Rapids, IA, USA; <sup>3</sup>College of Nursing, The University of Iowa, <sup>4</sup>Department of Emergency Medicine, The University of Iowa Hospitals and Clinics, Iowa City, IA, USA; <sup>5</sup>Midwestern University, Glendale, AZ, USA

**Background:** While the validity of pain assessment has been well documented, the underlying schema (ie, organized, preconceived ideas) of how individuals interpret numerical pain ratings is not well understood. This study's objectives were to examine numerical pain intensity ratings, from (0 to 10 cm on the visual analog scale [VAS]) across multiple severities of commonly experienced acute pain conditions to determine whether the ratings differed between these pain conditions and/or between individuals.

**Methods:** A community sample (N=365, 66% female) rated their anticipated pain intensity (VAS) for threshold, mild, moderate, severe, and tolerance level, using several common pain conditions: headache, toothache, joint injury, delayed-onset muscle soreness, burns, and "general pain."

**Results:** Cluster analysis revealed three subgroups of individuals, suggesting three types of underlying pain rating schema: 1) Low Rating subgroup (low VAS pain intensity ratings across all the pain severity categories); 2) Low/High Rating subgroup (low VAS pain intensity rating for mild, but high VAS pain intensity rating for severe pain); and 3) High Rating subgroup (high VAS pain intensity ratings across all the pain severity categories). Overall, differences between pain conditions were small: muscle soreness pain intensity was consistently rated lower than the other pain types across severities. The highest pain ratings varied between joint injury and general pain, depending on severity level. No effects of sex or current experience of pain were noted.

**Conclusion:** The results indicate that: 1) three distinct pain schemas were present in this community-based sample, indicating significant variation in how pain scales are utilized and/or interpreted between clusters of individuals; 2) pain ratings vary by condition, but these differences are minor; and 3) pain rating schemas are not significantly different between males and females or between individuals with and without current pain.

**Keywords:** VAS, muscle soreness, headache, toothache, joint pain

## Introduction

The psychometric assessment of pain has been well-studied and has contributed to our current understanding of the validity of subjective pain ratings, ratio scaling of pain, and the nonlinear power relationship in stimulus intensity pain rating curves.<sup>1-3</sup> However, the underlying cognitive appraisal strategies and schemas used when self-reporting pain intensity are not well understood. Schemas may be best described as organized, preconceived ideas or biases (ie, a conceptual framework) affecting how people attend to, interpret, and recall information. Accordingly, a pain rating schema is a conceptual framework used when communicating pain intensity as a numerical value.

Correspondence: Laura A Frey-Law  
Department of Physical Therapy and Rehabilitation Science, The University of Iowa, I-252 Medical Education Building, Iowa City, IA 52242-1190, USA  
Tel +1 319 335 9804  
Fax +1 319 335 9707  
Email [laura-freylaw@uiowa.edu](mailto:laura-freylaw@uiowa.edu)

Pain is an inherently subjective experience, involving sensory and emotional factors.<sup>4</sup> When asked to rate pain intensity, a cognitive determination of how to scale the experience must occur. While the endpoints of most numeric pain rating scales are predefined (eg, “no pain” versus “maximum experienced or imagined pain”), midpoint characterizations are often not operationally defined. As such, pain rating schemas may vary considerably between individuals (eg, what does a “3” out of 10 mean?). These variations may lead to miscommunication, misunderstandings, and compromised pain management, particularly if health care providers’ interpretations or schemas differ from patients’ schemas.<sup>5</sup> Clearly pain sensitivity varies between individuals;<sup>6–11</sup> however, it is not clear the degree to which underlying pain rating schemas differ between patients/individuals or between types of painful stimuli.

Cross-sectional comparisons have examined numerical and categorical (eg, mild, moderate, severe) pain ratings, particularly in terms of their association with pain interference in daily life, among patients with: amputation,<sup>12</sup> low back pain,<sup>13</sup> osteoarthritis,<sup>13,14</sup> diabetic peripheral neuropathy,<sup>15</sup> postoperative pain,<sup>16</sup> and cancer pain.<sup>17,18</sup> Collectively, these studies support the legitimacy of assigning numerical ratings to categorical severities (including those based on pain interference), despite some variations observed between studies and/or pain conditions (eg, mild pain is categorized as 1 to 3 in some studies, and 1 to 4 in others). Unfortunately, cross-sectional assessments of pain ratings at one severity or based solely on daily life interference due to pain do not provide insights into how pain is conceptualized across intensities or the types of pain for the same individual. Further, the experience of chronic pain may lead to an enmeshment of self- and illness-schemas,<sup>19</sup> thereby affecting an individual’s underlying pain schemas.

Our primary goal was to examine pain intensity ratings, using a 10 cm visual analog scale (VAS), for commonly experienced acute pain conditions (headache, toothache, joint pain, delayed-onset muscle soreness [DOMS], and burns), across multiple severities. We chose severities based on the widely used four-point verbal pain descriptor scale – none, mild, moderate, and severe – due to its clinical relevance (although “none” was not assessed in this study). In addition, pain threshold, the lowest intensity perceived as pain, and tolerance, the highest intensity of pain tolerable, were assessed due to their common use in experimental pain research. Rating pain associated with commonly experienced, acute pain conditions in a nonclinical sample allow for multiple measurements across severities and eliminate the inherent variability associated with clinical pathology in patient populations. We hypothesized that: 1) subgroups of individuals would emerge,

indicative of distinct underlying pain rating schema (ie, the cognitive appraisal and interpretation of numerical pain intensity across categorical severities); and 2) these schema would be situation-specific, as evidenced by differences in pain ratings between pain conditions (eg, DOMS could be rated lower than toothache pain). The secondary goals were to examine whether pain intensity ratings differed by sex or between people with and without existing pain.

## Methods

### Participants

A community sample (N=365, 66% female) was recruited through flyers and ads, with minimal exclusion criteria (ie, subjects had to be English-speaking, able to follow directions, and between 18 and 55 years) to obtain a sample that would be representative of the general adult population. All participants provided written informed consent prior to participation, as approved by the University of Iowa Biomedical Institutional Review Board, and were compensated for their time.

### Demographic and pain measures

The participants were first asked to complete a series of self-reported questions regarding their demographics (ie, age, race/ethnicity, sex, and general health status) and pain experiences: 1) past experience (yes/no) of five pain conditions (headache, muscle soreness, toothache, joint injury, and burn); 2) current pain (if any; rated using the 10 cm VAS); and 3) worst pain ever experienced (rated using the 10 cm VAS). To measure participants’ schemas for various pain conditions and severities, individuals were then instructed to answer 30 questions, using this instructional set for all items:

Please rate the following questions using the 10 cm visual analog scale (VAS), based on your personal experiences. Place a vertical mark on the line to show your answer, where far left indicates no pain and the far right indicates the maximum pain you have ever felt or can imagine feeling. Please be as honest as possible, there are no right or wrong answers.

The VAS ratings were measured in centimeters using a ruler.

A total of six pain conditions were assessed: pain “in general” and five specific acute pain experiences familiar to a large portion of the population<sup>20–26</sup> and thus likely to have been experienced by a majority of the sample: headache, toothache, joint pain (ankle sprain or knee pain), muscle soreness, and burns. Three common verbal descriptor severities

were assessed for each of the six pain conditions: mild, moderate, and severe. Self-reported threshold and tolerance were also assessed for a total of five categorical severity levels (see questions below for examples). Thus a total of 30 unique items (six pain conditions, rated across five severity levels) were assessed. Identical wording was used for each severity level across the 30 items, with the only between-item differences being the words for each specific pain condition (eg, joint injury versus toothache). Five examples, highlighting each severity level for the DOMS condition, are as follows:

- What would you rate as mild muscle soreness (1–2 days after activity)?
- What would you rate as moderate muscle soreness (1–2 days after activity)?
- What would you rate as severe muscle soreness (1–2 days after activity)?
- What is the lowest intensity you could perceive as muscle soreness (1–2 days after activity)?
- At what point would you rate muscle soreness (1–2 days after activity) as becoming intolerable?

Items in the pain questionnaire were randomized into two different orders (to minimize order or fatigue effects), with each order used for approximately half of the participants. Early versions were piloted on small cohorts of individuals and edited based on feedback and responses (not included in final analyses). A new questionnaire was developed for this study because there were no other measures that assess pain intensity across a variety of conditions and severities to evaluate pain rating schema.

## Statistical analyses

Descriptive statistics (mean, standard deviation [SD], median, and percentile) were calculated (SPSS v20; IBM Corporation, Armonk, NY, USA) for all 30 questionnaire items. Mean ratings were also calculated for each of the five pain severity levels, collapsing the scores across the six pain conditions, to obtain an average mild, moderate, severe, threshold, and tolerance score, as an estimate of how individuals generally rate acute pain for each severity level. Means and SD are reported in the text; means and standard error of means (SEM) are shown in all figures.

## Cluster analysis

Cluster analysis was used to assess for subgroups of individuals based on the three verbal descriptor pain severities (mild, moderate, and severe) for the six pain conditions (18 input variables). A two-step clustering algorithm, using

log-likelihood distance measures and Schwarz's Bayesian information criterion (BIC),<sup>27</sup> was used to determine the subgroups, automatically considering up to 15 clusters. This two-step clustering approach first assigns cases to "preclusters," then uses hierarchical clustering algorithms to determine the final group membership. The BIC provides a systematic and unbiased method to ascertain the optimal number of clusters.<sup>28,29</sup> Cluster membership was then saved as a categorical variable.

## Responses between cluster subgroups

To test whether pain ratings differed between cluster subgroups, pain conditions, or sex, a mixed repeated measures analysis of variance (ANOVA) (with the Huynh–Feldt correction) was used. The within-subjects variables included pain severity (five levels) and pain condition (six levels), and the between-subjects variables included cluster and sex. If significant between-group differences were observed, the effect sizes (Cohen's *d*) and post hoc follow-up tests were evaluated: Tukey's test for the between-subjects variables and Bonferroni corrections for the within-subjects variables, to adjust for multiple comparisons.<sup>30</sup> Next, ANOVAs were used to compare demographic variables between the clusters and to confirm differences in the mean pain rating variables between the final cluster subgroups. The absolute and relative differences in pain intensity ratings between each successive severity level (eg, mild versus moderate) were assessed and compared between clusters using ANOVA. Significance was set at  $\alpha = 0.05$  for all analyses.

## Minimal clinically important differences

Although minimal clinically important differences (MCIDs) for VAS pain ratings vary substantially based on patient population and statistical technique used, this may provide a useful criterion for comparing significant differences. Based on previous research, we used a midpoint value of 1.4 cm<sup>31</sup> as a MCID (reported range of 0.4–3.7 cm for a 10 cm scale),<sup>31–34</sup> as a general indicator of meaningful differences between groups and/or pain types in the current study.

## Questionnaire reliability and validity

To evaluate the reliability of our measure, short-term stability (2-week test–retest interval) was assessed in a subset of the study participants (the reliability cohort;  $n=49$ ). Multiple reliability statistics were considered to ensure consistency, including the Guttman split-half correlation coefficient; the Spearman–Brown coefficient; and intraclass correlation coefficient (ICC) (two-way fixed for consistency).<sup>35</sup>

The instrument error of the pain questionnaire was assessed as the mean square error (MSE) term for each categorical severity level (six in total) from the repeated ANOVA. Lastly, the internal consistency of the pain questionnaire was assessed using the Cronbach's alpha coefficient.

## Results

### Participants

A total of 365 participants were recruited, with an age range of 18 to 54 years and mean (SD) of 27.0 (10.6) years (Table 1). All participants were familiar with the majority of the pain conditions; all (100%) reported having previously experienced at least three of the five specific acute pain conditions considered, and 92.1% had experienced four or more of the conditions. Headache, muscle soreness, and burns were the most commonly (over 96%) experienced previously (Table 1). Toothaches were the least common, with approximately three-quarters (77%) of participants reporting previously experiencing a toothache. At the time of the questionnaire, nearly one-third (31.2%) of participants reported currently having pain, averaging 3.0 (SD =1.9) cm (range from 0.4 to 8.2 cm) on the 10 cm VAS. Muscle soreness, headache, and joint pain were the three most frequently cited

sources of current pain. The self-reported "maximum pain ever experienced" averaged 7.6 (SD =1.6) cm, with a range from 2.6 to 10.0 cm.

### Cluster analysis of subgroups

Results from the cluster analysis indicated individuals were classified approximately equally into three subgroupings (ie, clusters) of individuals (Figure 1 and Table 2). Due to missing data (four people did not answer the toothache items), only 361 of the 365 participants were included in the final cluster analyses. When considering one to 15 subgroups, three clusters produced the lowest BIC (4,108.7) and the largest ratio of distance measures (2.6), thus, the three-cluster solution was chosen over the two-cluster solution, for which BIC was 4200.6 and ratio of distance measures was 2.4. One cluster of individuals (n=105) rated VAS pain intensity consistently high across mild to severe pain (the High Rating subgroup). A second cluster (n=121) reported consistently low VAS pain ratings across the severities (the Low Rating subgroup). However, the third cluster (n=135) had a divergent pain rating pattern: these individuals assigned a low score to mild pain (similar to the Low Rating subgroup) but assigned a high score to the severe pain category (similar to the High Rating subgroup). Thus, we labeled this third cluster the Low/High Rating subgroup.

**Table 1** Participant characteristics

Variable	Range or category	Mean (SD) or n (% of sample)
Age	18 to 54	27.0±10.6
Sex	Males	121 (34.0%)
	Females	235 (66.0%)
Racial or ethnic heritage	White	312 (87.6%)
	Asian	21 (5.9%)
	Black/African American	8 (2.2%)
	Hispanic or Latino	7 (2.0%)
	More than one racial heritage	8 (2.2%)
General health	Excellent	154 (43.3%)
	Good	185 (52.0%)
	Fair	8 (2.2%)
	Poor	9 (2.5%)
Previously experienced <sup>†</sup>	Headache	349 (98.0%)
	Muscle soreness	346 (97.2%)
	Burn	345 (96.9%)
	Joint pain	308 (86.5%)
	Toothache	274 (77.0%)
Currently experiencing pain	Yes	111 (31.2%)
	Average VAS intensity (if in pain; 0–10 cm)	3.0 (1.9)

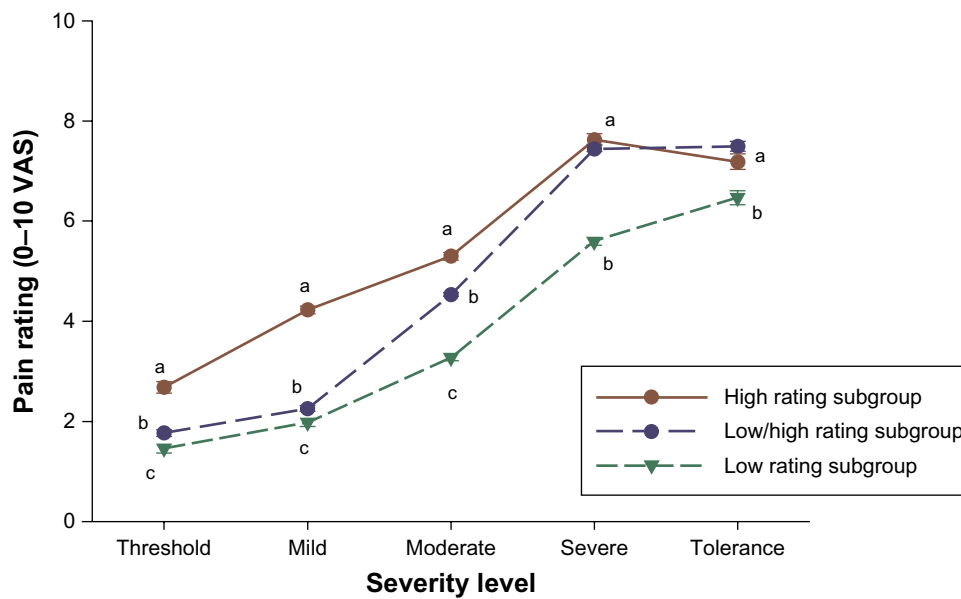
**Notes:** N=356 (nine participants did not complete the demographic questions). <sup>†</sup>Participants could endorse more than one previously experienced pain condition category.

**Abbreviations:** SD, standard deviation; VAS, visual analog scale.

### Pain rating differences

Pain ratings differed significantly between the categorical severity levels ( $F_{2,88,771.5}=1368.2$ ) ( $P<0.0001$ ) (Figure 2A), pain conditions ( $F_{4,42,1184.6}=49.2$ ) ( $P<0.0001$ ) (Figure 2B), and clusters ( $F_{2,268}=177.1$ ) ( $P<0.0001$ ) (Figures 1, 3, and 4). Sex differences were not significant ( $F_{1,268}=2.5$ ) ( $P=0.11$ ), and none of the sex interaction terms were significant ( $P>0.24$ ), indicating that men and women did not rate pain differently across the pain conditions or severity levels. Thus, the remaining results are reported for men and women collectively.

Box plots of the mean 10 cm VAS pain ratings by severity level (averaged across all six pain conditions) and by condition (averaged across all severity levels) are shown in Figure 2. The differences between the severities exceeded the MCID of 1.4 cm for all pairwise comparisons except for "threshold versus mild" (mean difference =0.74 cm) and "severe versus tolerance" (mean difference =0.23 cm). Although some of the mean differences between pain conditions were statistically significant, none reached our pre-determined MCID (range =0.02–0.997 cm).



**Figure 1** Overall mean (SEM) pain intensity ratings for each categorical severity level by cluster, averaging across all the pain conditions. **Notes:** Within each severity level, significant differences between the subgroups are indicated by different letters (a, b, and/or c), based on Tukey’s post hoc tests ( $P < 0.05$ ). The SEM estimates are very small, thus even small differences (eg, the threshold and mild pain ratings between the Low and Low/High subgroups) achieve significance. However the largest between-cluster differences occurred when rating mild and severe pain. **Abbreviations:** SEM, standard error of the mean; VAS, visual analog scale.

All of the two- and three-way interaction terms between pain condition, severity, and cluster were significant (Figures 3 and 4): cluster by pain condition ( $P < 0.0001$ ), cluster by severity ( $P < 0.0001$ ), severity by pain condition ( $P < 0.0001$ ), and severity by pain condition by cluster ( $P < 0.0001$ ). These significant interactions indicate that pain condition, categorical severity level, and cluster assignment all interact with one another to influence pain intensity ratings.

### Pain ratings between clusters

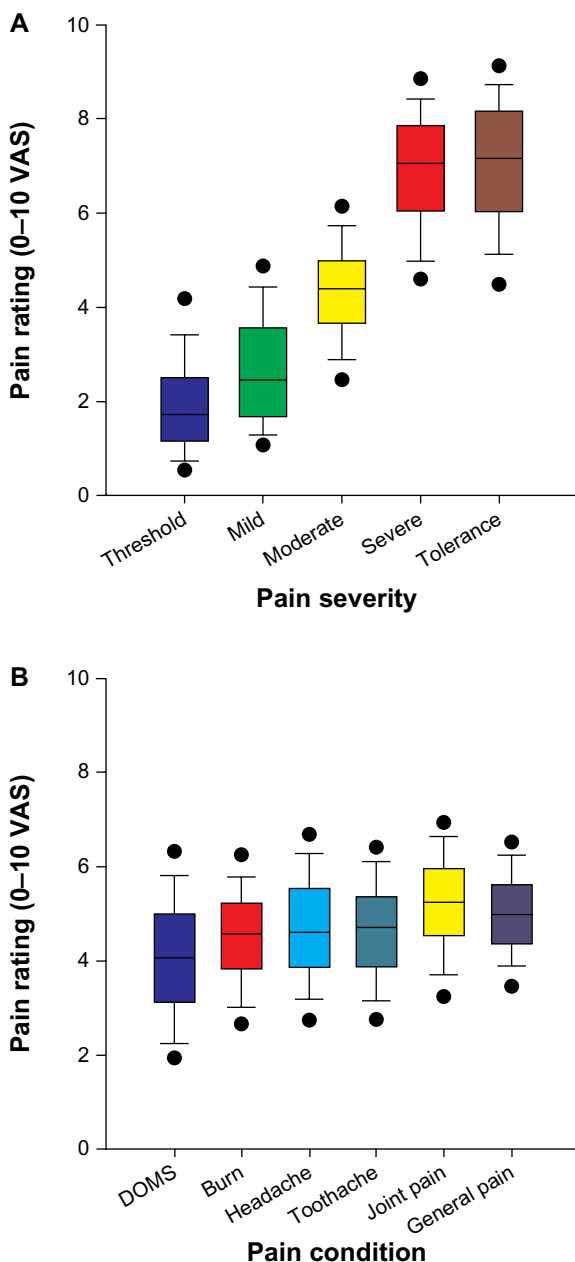
Confirming the cluster subgroupings, the post hoc tests revealed clear differences in the mean 10 cm VAS pain

ratings between clusters at each severity level, including threshold and tolerance (Figure 1). Individuals in the High Rating subgroup rated all pain severities significantly higher ( $P < 0.0001$ ) than did those in the Low Rating subgroup. The between-cluster differences ranged from 0.7/10 cm (for tolerance) to 2.3/10 cm (for mild pain); which corresponded to effect sizes that ranged from moderate to large,  $d = 0.5-2.7$ . Further, five of these pairwise differences achieved statistical significance and exceeded our criterion MCID of 1.4: both the Low Rating and Low/High Rating subgroups versus the High Rating subgroup for mild pain, the Low Rating versus the High

**Table 2** Cluster descriptive summaries and comparisons

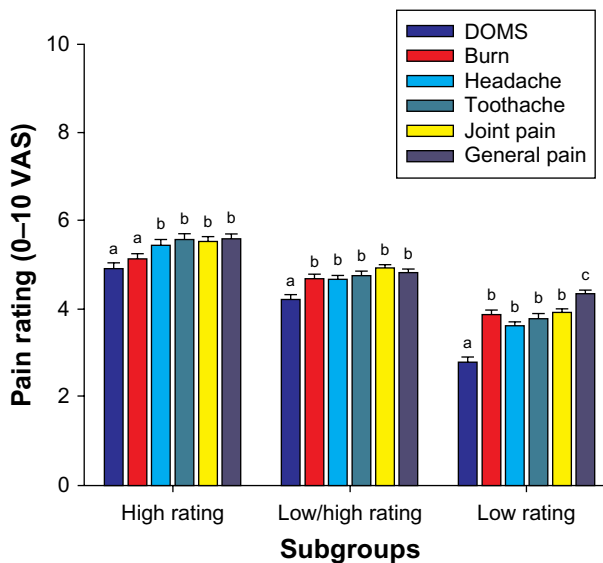
Variable	Cluster 1: high rating	Cluster 2: low/high rating	Cluster 3: low rating	P
n	105 (29.1%)	135 (37.4%)	119 (33.5%)	–
Age (years)	26.1 (10.7)	28.5 (10.8)	26.2 (9.4)	0.11
Males	33.6%	28.9%	41.2%	0.12
Current pain prevalence	32.7%	31.8%	29.4%	0.86
Current pain intensity, if in pain (0–10 cm VAS)	3.8 (1.9) <sup>a</sup>	3.3 (2.1) <sup>b</sup>	2.2 (1.5) <sup>c</sup>	0.001*
Maximum previous pain (0–10 cm VAS)	7.7 (1.6)	7.8 (1.5)	7.4 (1.6)	0.08
Mild–moderate pain difference ΔVAS	1.1 (1.1) <sup>a</sup>	2.4 (0.9) <sup>b</sup>	1.4 (0.9) <sup>c</sup>	<0.0001*
% increase	69.3 (79.7) <sup>a</sup>	212.0 (163.6) <sup>b</sup>	188.8 (181.1) <sup>b</sup>	<0.0001*
Moderate–severe pain difference ΔVAS	2.3 (1.1) <sup>a</sup>	2.9 (0.9) <sup>b</sup>	2.3 (0.9) <sup>a</sup>	<0.0001*
% increase	56.5 (38.4) <sup>a</sup>	77.2 (37.8) <sup>b</sup>	100.2 (69.4) <sup>c</sup>	<0.0001*

**Notes:** N=361. The values are expressed as n (%) or mean (SD). The percentages are based on the total number of individuals who responded to each question; One-way ANOVA was used for all continuous data; The Pearson chi-square test was used for dichotomous data. \*Significance was set at  $P \leq 0.05$ . <sup>a</sup>–<sup>c</sup>Means sharing a letter in their superscript are not significantly different from each other, based on the Tukey honestly significant difference post-hoc test. **Abbreviations:** ANOVA, analysis of variance; SD, standard deviation; VAS, visual analog scale.



**Figure 2** Box plots displaying the median, 25th to 75th percentiles (boxes), tenth to 90th percentiles (error bars), and fifth to 95th percentiles (dots) for the pain ratings, by (A) the categorical severity level (averaged across the pain conditions) and (B) the pain condition (averaged across the severity levels).  
**Abbreviations:** VAS, visual analog scale; DOMS, delayed-onset muscle soreness.

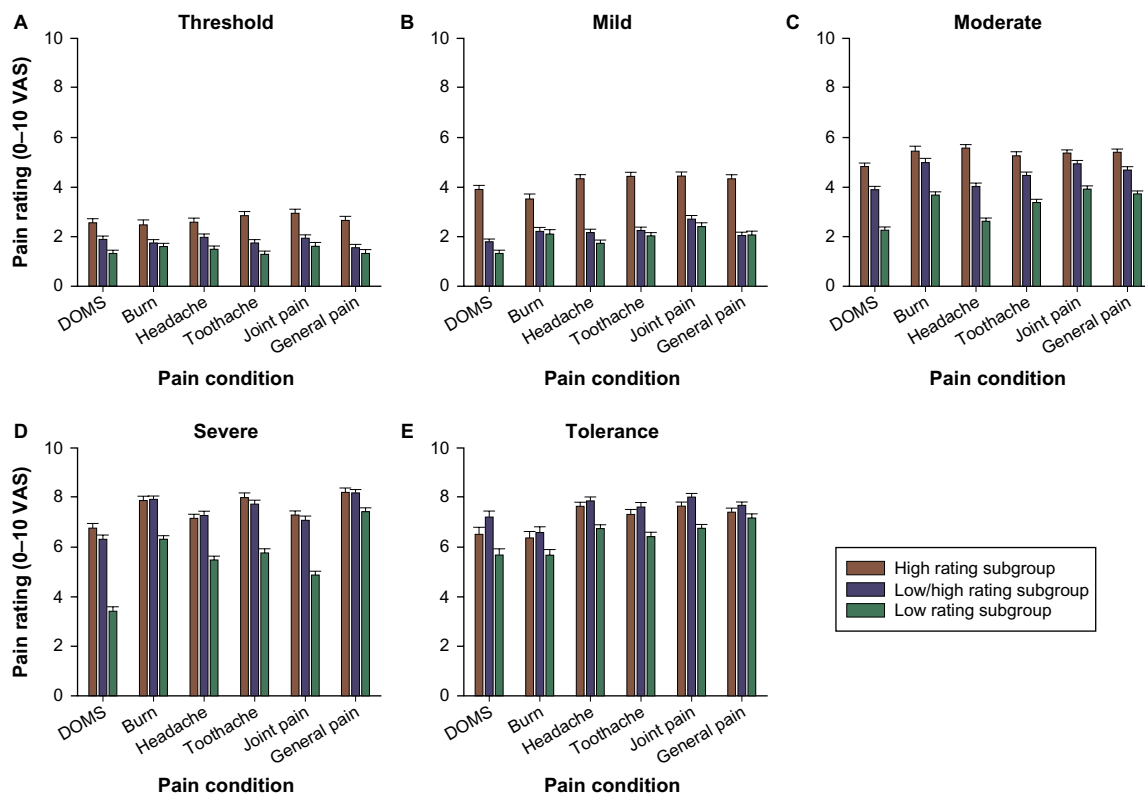
Rating subgroup for moderate pain, and both the Low/High Rating and the High Rating subgroups versus the Low Rating subgroup for severe pain (see Figure 1). When looking at the difference or slope in the numerical pain ratings from the mild to severe categories, the High Rating and Low Rating subgroups demonstrated remarkably similar approaches, with nearly parallel lines and a simple offset between them: the difference from mild to moderate pain was approximately half the difference



**Figure 3** Mean (SEM) pain intensity ratings for each pain condition averaged across severities (threshold, mild, moderate, severe, and tolerance), by cluster.  
**Notes:** Within each cluster, the significant between-pain condition differences (Tukey’s post hoc tests) ( $P < 0.05$ ) are indicated by different letters (a, b, and/or c). Note that DOMS was consistently rated lower than the other pain conditions, across all three clusters.  
**Abbreviations:** DOMS, delayed-onset muscle soreness; SEM, standard error of the mean; VAS, visual analog scale.

from moderate to severe pain, for both groups (Figure 1 and Table 2). Both the Low Rating and High Rating subgroups appeared to rely on a “restricted” range of available numerical pain rating intensities but at different ends of the spectrum. Whereas, the Low/High Rating subgroup rated pain similar to the Low Rating subgroup at the lowest intensities and the High Rating subgroup at the highest intensities, demonstrating a wider use of the full range of the numerical pain scale. Further, in the Low/High Rating subgroup, the increase in ratings from mild to moderate pain was approximately equal in magnitude to the increase from moderate to severe, which was different from what was observed for the other two clusters of individuals. Standardizing these increases in pain ratings did not significantly change these between-cluster differences (results not shown).

No differences in participant age, sex, proportion with current pain, or maximum pain ever experienced was found between the three clusters (Table 2). The only demographic difference observed between clusters was the reported pain intensity in those with current pain (31.2% of sample); individuals in the Low Rating cluster rated their current pain intensity significantly lower than those in either the High Rating ( $P = 0.002$ ) or the Low/High Rating subgroups ( $P = 0.05$ ) (Table 2).



**Figure 4** Mean (SEM) pain intensity rating for each pain condition by cluster, for the (A) threshold, (B) mild, (C) moderate, (D) severe, and (E) tolerance pain categories. **Notes:** The largest differences in pain ratings occur between the cluster subgroups and the severities, but relatively small differences are observed between the specific pain conditions. The significant differences were not indicated on this figure, for clarity.

**Abbreviations:** DOMS, delayed-onset muscle soreness; SEM, standard error of the mean; VAS, visual analog scale.

## Pain ratings between pain conditions

As hypothesized, muscle soreness pain intensity was rated significantly lower than were the other pain conditions within each severity level, with moderate to large effect sizes ranging from  $d=0.6$  to  $1.1$  ( $P<0.0001$ ) (Figures 3 and 4). Conversely, individuals rated general pain higher than all of the specific pain conditions, across the five categorical severities, but the effect sizes varied substantially – from small ( $d=0.2$ ) ( $P=0.01$ ) compared with joint pain, to large ( $d=1.1$ ) ( $P<0.0001$ ) compared with muscle soreness. This difference was most apparent in the Low Rating subgroup (see Figure 3). On average, differences in the anticipated pain ratings between the pain conditions were small (median effect size was  $d=0.3$ ) relative to the differences observed between severities or clusters, with most of the absolute 10 cm VAS differences falling

well below our criterion MCID of 1.4. The within-cluster post hoc comparisons demonstrated no significant differences between pain conditions at pain threshold ( $P>0.12$ ) (Figure 4A). However, across all the other severity levels (mild through tolerance), muscle soreness was consistently rated as the lowest or second lowest of all the pain conditions (Figure 4B–E). Further, the absolute VAS differences between muscle soreness and the other five pain conditions were the most likely to reach or exceed the MCID criterion of 1.4.

## Survey reliability and validity

The short-term (2 week, range 10 to 20 days) test–retest repeatability ( $n=49$ ) was good to excellent (Table 3). The Spearman–Brown coefficients, ICCs, between-session correlation coefficients, and Guttman split-half

**Table 3** Reliability statistics for the 30-item pain questionnaire (2-week test–retest interval)

Internal consistency		Test–retest measures			
Cronbach's alpha-first visit	Cronbach's alpha-second visit	Spearman–Brown coefficient	Intraclass correlation coefficient	Guttman split-half coefficient	Pearson's correlation coefficient
0.96	0.96	0.90	0.98	0.90	0.82

coefficients demonstrated good-to-high test–retest reliability (0.82 to 0.98). The 30-item instrument was internally consistent at each test session, with alpha coefficients at 0.96. The construct validity of the instrument was indicated by the clear ordering of the mean pain ratings, where threshold < mild < moderate < severe (see Figure 1).

The instrument error estimates, based on the 2-week test–retest SEM, ranged from 0.40 to 0.78 across the severities. Accordingly, any observed differences between the pain ratings (eg, between conditions or severities) that were less than these SEM values may have been due to error alone. Further, these error estimates were one-third to one-half of the magnitude of the MCID we used for comparison purposes. Thus any meaningful differences observed were well beyond the instrument error levels.

## Discussion

This was a novel study that examined anticipated numerical pain intensities across a range of categorical severities for commonly experienced pain conditions. Our use of diverse pain experiences was particularly innovative and important because prior research suggests that pain sensitivity varies between different types of pain.<sup>36</sup> The results indicate that: 1) significant variation occurs in how numerical pain ratings are utilized/interpreted (ie, pain schemas) between subgroupings of individuals for a given categorical severity; 2) pain ratings can vary by pain condition, but these differences are small relative to the differences observed between subgroups of individuals; 3) pain rating schema are not significantly different between males and females; and 4) existing pain does not have a significant influence on acute pain rating schemas. Altogether, these findings are potentially important for clinical and research settings because they indicate that individuals may differ in their underlying cognitive appraisal strategies when self-reporting numeric pain intensity, for mild, moderate, and severe pain, and also, when considering pain threshold and tolerance.

The participants were asked to consider several common acute pain experiences that tend to vary in their pain quality (ie, sharp, dull, or aching) and anatomical source (ie, joints, muscles, skin, head, or tooth pain). Overall, similar pain rating schemas were observed across these pain conditions, with any significant differences between the pain conditions being smaller than the differences occurring between clusters of individuals. Thus, underlying pain schemas or cognitive assessments of pain appear to be relatively consistent across pain types, despite known differences in transmission and processing between various sources of pain.<sup>37,38</sup>

The three clusters of subjects observed in this study differed on their ratings of pain severity levels (ie, mild, moderate, and severe), but did not differ in demographic (sex or age) or pain-related variables (prior pain experiences or existing pain). These three clusters appear to differentiate three underlying pain rating schema: 1) the Low Rating subgroup demonstrated a consistent bias to numerically rate pain relatively low across all severities; 2) the High Rating subgroup demonstrated a consistent bias to numerically rate all categorical severities on the higher side of the numerical scale; and 3) the Low/High Rating subgroup used the numerical rating scale in its entirety, with low VAS ratings for the lowest categorical severities and high VAS ratings for the highest categorical severities. These results may impact our understanding of individual differences in pain sensitivity. That is, heterogeneity in pain sensitivity observed in prior studies<sup>6,8,10,39,40</sup> may be partially attributable to underlying differences in pain rating schemas.

The ranges of VAS intensities reported by our community sample for mild, moderate, and severe pain are consistent with previous cross-sectional studies in patient populations, some of which were based on pain interference scores. For example, mild pain has typically encompassed ratings from 1 to 3,<sup>41–43</sup> or 1 to 4,<sup>5,13,17,18,39,44–46</sup> which is similar to the fifth to 95th percentile range of 1.1 to 5.0 in our cohort, despite the between-study differences in collection and statistical methodologies.

However, our results would suggest that single “cutpoints” to differentiate numerical pain ratings into categorical levels may not adequately reflect the individual differences associated with decision-making and the underlying cognitive approach used to evaluate the numerical pain scales. Indeed, approximately one-third of the study cohort (ie, the High Rating subgroup) rated mildly painful conditions as nearly twice as painful on the 10 cm VAS as the remaining two-thirds of individuals (Low Rating and Low/High Rating subgroups). Further, this High Rating subgroup rated mild severities higher than or equal to the moderate severities rated by the other two subgroups. Thus, clinicians or researchers who assign cutpoints to numerical pain ratings to translate them into mild, moderate, or severe severity categories may misinterpret an individual’s pain experience.

We did not see a significant difference between males and females in pain rating schema. This is consistent with a recent review indicating little to no sex differences in several assessments of pain sensitivity.<sup>47</sup> Our findings also suggest that previously observed sex differences in pain tolerance, ie, for thermal<sup>48</sup> and cold pressor pain,<sup>7,49</sup> are not simply due to a



difference in pain rating schema – sex differences in tolerance are more likely a result of underlying physiological differences (eg, in subcutaneous fat layers) or social–behavioral differences, where men may be more reluctant to “quit” a pain tolerance task. Indeed, gender roles and expectations have accounted for the sex differences in pain ratings in prior studies.<sup>50,51</sup>

The most notable difference between the clusters was how individuals rated their existing pain intensity: Individuals in the Low Rating cluster rated “current” pain the lowest of all three clusters. This is consistent with their observed pain cluster schema, which was to use the lowest half of the numerical range available. However, it is impossible to differentiate whether these differences in existing pain were due to: 1) less noxious painful stimuli; 2) lower pain sensitivity; 3) their underlying pain rating schema used to communicate their pain; or 4) some combination of these three.

Several differences between pain conditions were also observed, largely following the initial hypothesis that muscle soreness would be rated lowest on the 10 cm VAS for each severity. It is likely that the affective experience (eg, fear or anxiety) or expectations associated with each condition differ. For example, DOMS, a common nonpathological condition that results from unaccustomed exercise, is sometimes considered to be a “good pain.”<sup>52</sup> Thus, pain-related fear, catastrophizing, and/or anxiety may be more likely to occur with presumably more threatening clinical or injurious conditions, such as headache, toothache, or joint pain, than with DOMS. Prior research has demonstrated that the affective dimension of pain can modify pain schema.<sup>53</sup> Accordingly, the underlying schema developed based on prior painful experiences is likely to have been influenced by the affective dimensions associated with each acute pain condition. Although we did not measure these factors in this study, they may be important to assess in future research.

Several study limitations are worth noting. Our study utilized a community sample of individuals, with and without pain, rather than a patient population. While nearly one-third reported experiencing pain at the time of the study, it is unknown what proportion had chronic pain. We recognize that the presence of chronic pain may further influence pain schemas through an enmeshment of the self- and illness schemas.<sup>19</sup> The study cohort was also predominantly young and Caucasian, and the generalizability of these results to older adults or other racial or ethnic groups is unknown. However, previous studies suggest that pain rating boundaries do not significantly differ across demographic groups.<sup>5,44</sup> Future studies are needed to better determine whether additional participant characteristics, such as age, personality,

pain-related catastrophizing or fear, past pain experiences, chronic pain, and/or social or cultural environments predict cluster membership. Lastly, because this is the first questionnaire to specifically assess pain rating schema through pain intensity assessments across severities for multiple pain conditions, we were unable to directly assess the convergent or divergent validity of our instrument.

In summary, these results provide insights into the underlying variations in the conceptualization of numerical pain intensity across categorical severities, ie, pain rating schema. We found that three general clusters emerged, indicating that three distinct approaches were used to rate pain across a range of conditions and severities. Accordingly, it may be valuable to assess both numerical intensity and categorical severity to better understand how each individual is evaluating their pain experience. For example, this study demonstrates that a pain rating of “4” could be perceived as mild, moderate, or even severe pain, depending on the individual’s underlying pain rating schema. Thus, we recommend that clinicians and researchers consider asking both intensity ratings as well as categorical interpretations (ie, “Can you rate your pain on a 0 to 10 scale?” and “Would you classify your pain as mild, moderate, or severe?”). Lastly, subgrouping individuals based on pain schema may benefit pain heterogeneity studies examining the predictors of pain sensitivity or pain phenotypes to better differentiate cognitive assessment from other sources of heterogeneity (eg, physiological or neurochemical). Ultimately, this information may be used to enhance communication between patients and their care providers and allow for improved comparisons across research studies.

## Author contributions

LAF, MM, and AMW conceived and designed the experiments, with input from JEL. LAF, AMW, and JEL analyzed the data. All authors discussed the results and commented on the manuscript. LAF and JEL wrote the initial manuscript draft. All four authors actively reviewed and revised the manuscript, and approved the final submitted manuscript.

## Acknowledgments

This work was supported in part by funding from The University of Iowa Institutional Research Experience for Undergraduates program (MM) and the Medical Student Summer Research Fellowship program (AMW). In addition, training grants from the National Institutes for Health (NIH) supported authors during portions of this study: LAF (K01 AR056134 and K12 HD055931) and JEL (Interdisciplinary Pain Research Program, T32 NS045549).

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Gracely RH, McGrath F, Dubner R. Ratio scales of sensory and affective verbal pain descriptors. *Pain*. 1978;5(1):5–18.
- Price DD, Bush FM, Long S, Harkins SW. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain*. 1994;56(2):217–226.
- Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*. 1983;17(1):45–56.
- Merskey H, Bogduk N. A current list with definitions and notes on usage. In: Merskey H, Bogduk N, editors. *Classification of Chronic Pain*. 2nd ed. Seattle, WA: IASP Press; 1994:209–214.
- Li KK, Harris K, Hadi S, Chow E. What should be the optimal cut points for mild, moderate, and severe pain? *J Palliat Med*. 2007;10(6):1338–1346.
- Lee JE, Watson D, Frey-Law LA. Psychological factors predict local and referred experimental muscle pain: a cluster analysis in healthy adults. *Eur J Pain*. 2013;17(6):903–915.
- Dixon KE, Thorn BE, Ward LC. An evaluation of sex differences in psychological and physiological responses to experimentally-induced pain: a path analytic description. *Pain*. 2004;112(1–2):188–196.
- Frey Law LA, Sluka KA, McMullen T, Lee J, Arendt-Nielsen L, Graven-Nielsen T. Acidic buffer induced muscle pain evokes referred pain and mechanical hyperalgesia in humans. *Pain*. 2008;140(2):254–264.
- George SZ, Dannecker EA, Robinson ME. Fear of pain, not pain catastrophizing, predicts acute pain intensity, but neither factor predicts tolerance or blood pressure reactivity: an experimental investigation in pain-free individuals. *Eur J Pain*. 2006;10(5):457–465.
- Lee JE, Watson D, Frey Law LA. Lower-order pain-related constructs are more predictive of cold pressor pain ratings than higher-order personality traits. *J Pain*. 2010;11(7):681–691.
- Thorn BE, Clements KL, Ward LC, et al. Personality factors in the explanation of sex differences in pain catastrophizing and response to experimental pain. *Clin J Pain*. 2004;20(5):275–282.
- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. *J Consult Clin Psychol*. 2001;69(4):655–662.
- Zelman DC, Hoffman DL, Seinfeldin R, Dukes EM. Development of a metric for a day of manageable pain control: derivation of pain severity cut-points for low back pain and osteoarthritis. *Pain*. 2003;106(1–2):35–42.
- Kapstad H, Hanestad BR, Langeland N, Rustøen T, Stavem K. Cutpoints for mild, moderate and severe pain in patients with osteoarthritis of the hip or knee ready for joint replacement surgery. *BMC Musculoskelet Disord*. 2008;9:55.
- Zelman DC, Dukes E, Brandenburg N, Bostrom A, Gore M. Identification of cut-points for mild, moderate and severe pain due to diabetic peripheral neuropathy. *Pain*. 2005;115(1–2):29–36.
- Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain*. 1997;72(1–2):95–97.
- Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*. 1995;61(2):277–284.
- Paul SM, Zelman DC, Smith M, Miaskowski C. Categorizing the severity of cancer pain: further exploration of the establishment of cutpoints. *Pain*. 2005;113(1–2):37–44.
- Pincus T, Morley S. Cognitive-processing bias in chronic pain: a review and integration. *Psychol Bull*. 2001;127(5):599–617.
- Rasmussen BK, Jensen R, Schroll M, Olesen J. Epidemiology of headache in a general population – a prevalence study. *J Clin Epidemiol*. 1991;44(11):1147–1157.
- Gran JT. The epidemiology of chronic generalized musculoskeletal pain. *Best Pract Res Clin Rheumatol*. 2003;17(4):547–561.
- Hasselström J, Liu-Palmgren J, Rasjö-Wrååk G. Prevalence of pain in general practice. *Eur J Pain*. 2002;6(5):375–385.
- Linton SJ. New research provides new avenues for prevention: an overview of the book. In: Linton SJ, editor. *New Avenues for the Prevention of Chronic Musculoskeletal Pain. Vol 12, Pain Research and Clinical Management*. Amsterdam: Elsevier; 2002;12:1–6.
- Stovner Lj, Hagen K, Jensen R, et al. The global burden of headache: a documentation of headache prevalence and disability worldwide. *Cephalgia*. 2007;27(3):193–210.
- Bastos JL, Gigante DP, Peres KG. Toothache prevalence and associated factors: a population based study in southern Brazil. *Oral Dis*. 2008;14(4):320–326.
- Cohen LA, Bonito AJ, Akin DR, et al. Toothache pain: behavioral impact and self-care strategies. *Spec Care Dentist*. 2009;29(2):85–95.
- Schwarz G. Estimating the dimension of a model. *Ann Stat*. 1978;6(2):461–464.
- Fraley C, Raferty AE. How many clusters? Which clustering method? Answers via model-based cluster analysis. *Computer Journal*. 1998;41(8):578–588.
- Kayri M. Two-step clustering analysis in researches: A case study. *EJER*. 2007;28:89–99.
- Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 2000.
- Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. *J Shoulder Elbow Surg*. 2009;18(6):927–932.
- Stauffer ME, Taylor SD, Watson DJ, Peloso PM, Morrison A. Definition of nonresponse to analgesic treatment of arthritic pain: an analytical literature review of the smallest detectable difference, the minimal detectable change, and the minimal clinically important difference on the pain visual analog scale. *Int J Inflam*. 2011;2011:231926.
- Lauche R, Langhorst J, Dobos GJ, Cramer H. Clinically meaningful differences in pain, disability and quality of life for chronic nonspecific neck pain – a reanalysis of 4 randomized controlled trials of cupping therapy. *Complement Ther Med*. 2013;21(4):342–347.
- Auffinger BM, Lall RR, Dahdaleh NS, et al. Measuring surgical outcomes in cervical spondylotic myelopathy patients undergoing anterior cervical discectomy and fusion: assessment of minimum clinically important difference. *PLoS One*. 2013;8(6):e67408.
- Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231–240.
- Hastie BA, Riley JL, Robinson ME, et al. Cluster analysis of multiple experimental pain modalities. *Pain*. 2005;116(3):227–237.
- Wall PD, Woolf CJ. Muscle but not cutaneous C-afferent input produces prolonged increases in the excitability of the flexion reflex in the rat. *J Physiol*. 1984;356:443–458.
- Takiguchi SA, Myers SA, Yu M, Levy MM, McNamara JJ. Clinical and financial outcomes of lateral rotation low air-loss therapy in patients in the intensive care unit. *Heart Lung*. 1995;24(4):315–320.
- Chow E, Doyle M, Li K, et al. Mild, moderate, or severe pain categorized by patients with cancer with bone metastases. *J Palliat Med*. 2006;9(4):850–854.
- Dannecker EA, George SZ, Robinson ME. Influence and stability of pain scale anchors for an investigation of cold pressor pain tolerance. *J Pain*. 2007;8(6):476–482.
- Dihle A, Helseth S, Paul SM, Miaskowski C. The exploration of the establishment of cutpoints to categorize the severity of acute postoperative pain. *Clin J Pain*. 2006;22(7):617–624.
- Hanley MA, Masedo A, Jensen MP, Cardenas D, Turner JA. Pain interference in persons with spinal cord injury: classification of mild, moderate, and severe pain. *J Pain*. 2006;7(2):129–133.
- Mendoza TR, Chen C, Brugger A, et al. Lessons learned from a multiple-dose post-operative analgesic trial. *Pain*. 2004;109(1–2):103–109.

44. Palos GR, Mendoza TR, Mobley GM, Cantor SB, Cleeland CS. Asking the community about cutpoints used to describe mild, moderate, and severe pain. *J Pain*. 2006;7(1):49–56.
45. Jensen MP, Smith DG, Ehde DM, Robinsin LR. Pain site and the effects of amputation pain: further clarification of the meaning of mild, moderate, and severe pain. *Pain*. 2001;91(3):317–322.
46. Turner JA, Franklin G, Heagerty PJ, et al. The association between pain and disability. *Pain*. 2004;112(3):307–314.
47. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and pain perception – part 2: do biopsychosocial factors alter pain sensitivity differently in women and men? *Pain*. 2012;153(3):619–635.
48. Fillingim RB, Maixner W, Kincaid S, Silva S. Sex differences in temporal summation but not sensory-discriminative processing of thermal pain. *Pain*. 1998;75(1):121–127.
49. Cross T, Thorn B, Ward L. Pain and depression: Relationships of pain catastrophizing, depression, and dysfunctional attitudes. *J Pain*. 2004;5(Suppl 3):S99.
50. Buckingham RM, Charles MA, Beh HC. Extraversion and neuroticism, partially independent dimensions? *Pers Individ Diff*. 2001;31(5):769–777.
51. Klein CS, Allman BL, Marsh GD, Rice CL. Muscle size, strength, and bone geometry in the upper limbs of young and old men. *J Gerontol A Biol Sci Med Sci*. 2002;57(7):M455–M459.
52. McFarland EG, Compton SP, Dawson CA. Guide to good and bad pain for the health/fitness professional. *ACSMs Health Fit J*. 2003;7(4):11–16.
53. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365–376.

### Journal of Pain Research

## Publish your work in this journal

The Journal of Pain Research is an international, peer-reviewed, open access, online journal that welcomes laboratory and clinical findings in the fields of pain research and the prevention and management of pain. Original research, reviews, symposium reports, hypothesis formation and commentaries are all considered for publication.

Submit your manuscript here: <http://www.dovepress.com/journal-of-pain-research-journal>

Dovepress

The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.