Detection-Theory Analysis of Scaling and Discrimination Tasks: Responses to Noxious Thermal Stimuli.

CHEE-WEE TAN

Queen Margaret University, Edinburgh, United Kingdom

SHEA T. PALMER

University of the West of England, Bristol, United Kingdom

DENIS J. MARTIN

Teesside University, Middlesborough, United Kingdom

PATRICIA A. ROCHE

Queen Margaret University, Edinburgh, United Kingdom

Contact person:
Chee-Wee Tan

Address:
Queen Margaret University
Duke Street, Edinburgh
EH6 8HF
United Kingdom.

Telephone number:
(44) 131 3173641

Email:
ctan@qmu.ac.uk

Abstract

This study's main purpose was to examine the sensitivity estimates obtained from scaling and discrimination approaches for nociception assessment in healthy individuals. This investigation may inform future applications in diagnostic procedures for painful conditions. The models of psychophysical judgment based on Durlach and Braida (1969), Laming (1984) and Irwin et al. (1991) were used as the common analytical framework. Noxious thermal contact stimuli were used. The results showed that the scaling approach produced lower detection theory sensitivity estimates compared to the discrimination approach. The additional judgment variance in scaling tasks could explain this lowered sensitivity. The relative judgmental variance value of 2.18 in this study is lower than previous investigations. This was probably due to the present study employing a relatively smaller stimulus range. The authors propose that the theoretical framework in this study may be used in future studies to investigate the different dimensions of pain perception.

Author's note

We would like to thank Michael Hautus, Gary Rollman and all the reviewers for their invaluable comments on the manuscript.

Introduction

Psychophysical measurements are commonly used to quantify the judgments of noxious stimuli. There are two main psychophysical approaches in pain research. The first, the direct scaling method, involves the participant estimating their perception of the intensity of the noxious stimulus. An example of such an approach using a bounded scale is the Visual Analogue Scale (VAS) (Price, 1994). The second, the discrimination method, involves discriminating between two stimuli of different noxious intensities. An example is the yes-no experiment, where the participant is required to state whether the stronger or more noxious stimulus was presented. Both direct scaling and discrimination methods are used frequently in pain research, for example in brain imaging studies (Pertovaara et al., 2004), neurophysiological studies (Nahra & Plaghki, 2005), and clinical studies (Kemperman et al., 1997). The topic of interest for our research group is the potential use of psychophysical methods as differential diagnostic tools for painful conditions. For example, a clinical scenario may consist of two similar patient groups within a broader patient grouping manifesting different responses to scaling or discrimination methods. This may provide the basis for further investigations determining if the differences may involve biological, affective or cognitive dissimilarities (Petersen & Rowbotham, 2006). In order to proceed with this program, the comparison of responses using scaling and discrimination psychophysical methods by healthy individuals, intended for clinical application, needs to be compared. However, critics have argued that the discrimination method may not quantify the perceived noxiousness of the stimuli, rather it only compares the perceived intensity magnitudes of the stimuli

(Rollman, 1977; Craig & Rollman, 1999). Based on this argument, discrimination methods may lack construct validity in measuring the perceived noxiousness of stimuli.

To investigate whether discrimination methods do measure the perceived noxiousness of the stimuli and that this is comparable to the direct scaling methods, Irwin & Whitehead (1991) and Irwin, Hautus, Dawson, Welch, & Bayly (1994) have used signal detection theory as a common analytical framework for the data obtained for both methods. The index used for estimating the sensitivity was d'. The theories behind the analytical framework were originally proposed by Braida & Durlach (1972) and Laming (1984, 1997) for comparing different psychophysical methods in sensory perception.

Irwin and coworkers found that the psychometric function obtained for the discrimination method was similar to the direct scaling method. This provided some evidence to support the use of discrimination in the measurement of noxious stimuli perception. However, the sensitivities obtained for the direct scaling method were reduced compared to the discrimination method. This may be partially explained by the way judgments are made for both methods. Laming (1984) argued that the assigned magnitude to physical stimuli by participants' may be considered to be only nominal. Therefore information obtainable from these judgments is limited to the response frequencies assigned to different stimuli. Using the information from the response frequencies, the discriminability between the stimuli may be estimated. From this perspective, the sensitivity estimate obtained through sensory judgment is a rough estimate of sensory discrimination. According to Braida & Durlach's theory (1972), the reduction in d' obtained for the direct scaling method was due to the way participants judged the stimuli presented during the task. For the direct scaling method, the participant compared the sensation of the stimuli with the range of stimuli presented within the context of the experiment (Braida & Durlach, 1972). In addition to that, the responses made by participants for the direct scaling method may be autocorrelated to the previous response (Laming, 1984). These potential sources of judgment 'variance' may degrade the *d′* for the direct scaling method. These judgment variances may not be present, or at least minimized, in the discrimination method.

Braida & Durlach's theory (1972) states that it is possible to estimate the relative extent of the judgment variance inherent in the direct scaling method compared to the variance associated with the discrimination method. According to the standard model of signal detection theory, the sensitivity is $d'_D = (\mu_2 - \mu_1) / \sigma_D$, where d'_D is the discriminability between the two adjacent classes of stimuli, μ_1 and μ_2 are the means of the normal probability densities and σ_D is their common standard deviation. When the standard model is extended to encompass the additional variance inherent in the direct scaling method, then $d'_S = (\mu_2 - \mu_1)/(\sigma_D^2 + \sigma_S^2)^{1/2}$, where d'_S is the discriminability between the two adjacent classes of stimuli in the direct scaling method, σ_D^2 is the stimulus variance associated with the discrimination method, and σ_S^2 is the judgmental variance associated with the direct scaling method (Macmillan & Creelman, 2005, p134). An estimate of the additional judgmental variance relative to the stimulus variance can be obtained by

$$\sigma_S^2 / \sigma_D^2 = \left(d'_D / d'_S\right)^2 - 1 \ , \qquad (1)$$

(Durlach & Braida, 1969; Macmillan & Creelman, 2005).

One of the aims of the present investigation was to extend the analytical framework proposed by Irwin et al. (1994) to psychophysical measurements of responses from noxious thermal stimuli. The relevance of noxious thermal stimulus in pain research is established at several levels. It is the one of the most commonly used physical stimulus used for evoking experimental pain (Gracely, 2005). Neurobiologically, thermal stimulus activates a known narrow range of primary afferent fiber nociceptors, namely C-fiber and Type I and Type II A-fiber nociceptors (Meyer & Campbell, 1981; Treede, Meyer, Raja & Campbell, 1995). At the molecular level, noxious thermal stimulus activates a non-selective cation channel, the transient receptor potential vanilloid-1 (TRPV1) receptor, which is a potential therapeutic target for pharmacological management of pain (Caterina, et al., 1997).

Another aim was to verify previous findings by Braida & Durlach (1972), Irwin & Whitehead (1991) and Irwin et al. (1994) that the direct scaling method produced decreased d' compared to the discrimination method. This study will also estimate the relative amounts of judgment variance between both methods.

Methods

Participants

Participants were recruited from the students and staff of Queen Margaret University Edinburgh (QMU) using convenience sampling. Six healthy volunteers (4 women and 2 men) took part in the experiment. The median age of the participants was 28 years (range: 21-35 years).

Ethical approval

This study was approved by Queen Margaret University's Research Ethics Committee. All the participants provided written informed consent for participation in this experiment.

Inclusion and exclusion criteria

The inclusion criteria were: a) any healthy individual 18 years or over, b) able to provide consent for participation in this study. The exclusion criteria were: a) the presence of medical conditions that caused anesthesia to the tested limb, or the consumption or application of medication that caused analgesia or anesthesia on the tested limb, b) any wounds or injury to the tested limb.

Apparatus and Stimuli

The thermal stimuli were applied on the ventral surface of both forearms. A Thermotest (Somedic AB, Sweden) was used to administer heat stimuli via a contact thermode (with surface measuring 25mm × 50mm). Heating and cooling of the contact surface were achieved through a Peltier element housed within the thermode. The stimulus sets (45°C, 46°C, 47°C and 48°C) were pre-programmed using the EXPOSURE software (Somedic AB, Sweden).

Procedure

There were two tasks: a magnitude description task (MDT), representing the direct scaling method, and an intensity resolution task (IRT), representing the discrimination method. Each task was performed on different forearms for each participant, chosen at random without replacement. All randomizations within this experiment were performed using an online randomization plan generator (http://www.randomization.com). The participant completed both the MDT and IRT tasks within the same day. Twenty practice trials, similar to the actual trials, were presented at the beginning of every task for familiarization.

Magnitude description and intensity resolution tasks

The one-interval rating task was used for both IRT and MDT. Each trial began with the experimenter instructing the participant to place his/her forearm on the thermode (pre-set at the relevant testing temperature). A trial contained an observation period of 3 seconds. An automated auditory signal indicated to the participant to remove his/her forearm from the thermode. If the participant was not able to tolerate the full length of stimulus application, he/she was allowed to lift their forearm away from the thermode, although no participants did so during the study. There was an interstimulus interval (ISI) of 10 seconds before the next trial started.

The stimulus set for both tasks consisted of 4 temperatures: 45°C, 46°C, 47°C and 48°C. For the IRT trials, each trial presented one of two temperatures. There was equal probability of presentation for either of the two temperatures. There were three stimulus pairs in total: 45°C and 46°C, 46°C and 47°C, and 47°C and 48°C. The stimulus pair presentation was randomized. The three stimulus pairs of the IRT clocked a total of 240 trials per participant (80 trials for each of the three stimulus pairs). For the MDT trials, each trial presented one of four temperatures. Again, there was equal probability of presentation for any one of the four temperatures. There were a total of 160 trials per participant clocked for the MDT (40 trials for each of the four temperatures). The orders of trial presentation for both tasks were randomized.

The participants verbally indicated their judgments to the experimenter and these were recorded. For both tasks, responses were made based on response sets with six categories. The magnitude description task (MDT) required the participant to estimate the perceived magnitude of the stimulus presented based on a response set with six descriptions of sensory quality. The six categories are, in increasing magnitude: warm, hot, faint pain, painful, very painful and severe pain (Figure 1A). For the intensity resolution task (IRT), the participant rated their degree of confidence on whether the stimulus presented was the higher or lower intensity of a pair of stimulus intensities (Figure 1B).

(Figure 1 about here)

The participants were told the temperature of the administered stimulus at the end of each trial, i.e. trial-by-trial feedback was provided for both tasks. Feedback was introduced to minimize bias in participants' judgments through the comparison of observations with a weighted average of stimulus effects, also known as the adaptation level effect (Helson, 1964). In a pilot study, it was identified that participants were apprehensive in making judgments when no feedback was given. Feedback reassured the participants when making judgments and encouraged the use of the entire response set.

For the MDT, the participants received the following instructions:

In this experiment you will be asked to judge the intensities of heat stimuli presented to you. The judgment method involves assigning categories with descriptions to match the intensities of the heat sensations you will experience (Fig. 1A shown to the participant). There are six categories of intensities. Verbally indicate to the experimenter the category number with a description that matches most closely to the sensation you experienced. After you have done this, you will be told the temperature of the heat stimulus just presented to you.

For the IRT, the participants received the following instructions:

In this experiment you will be asked to determine which one of two heat stimuli was presented to you. One stimulus is hotter than the other. Your task is to indicate whether the presented stimulus was the higher or the lower intensity and how confident you are in making that decision. There are six categories to describe your decision (Fig. 1B shown to the participant). Verbally indicate to the experimenter the category number with a description that matches most closely to your decision. After you have done this, you will be told the temperature of the heat stimulus just presented to you.

Prevention of hyperalgesia, heat injury and wind-up

Two specific procedures were implemented to prevent hyperalgesia onset and heat injury of the test sites. The first procedure involved the participant being instructed to shift the position of the thermode to an adjacent forearm skin area at the beginning of a new trial. The second procedure involved the enforcement of an interstimulus interval (ISI) of 10 seconds. This latter procedure also minimized the effect of preceding stimuli increasing the perceived noxiousness of latter trials as a result of temporal summation. This phenomenon of noxious temporal summation is termed wind-up (Price, Hu, Dubner & Gracely (1977); Staud, Price, Robinson, Mauderli & Vierck, 2004). Wind-up is usually maintained when the ISI is less than 3 seconds. The participant's forearm was checked by the experimenter for signs of heat injury after every 20 trials or if there was a concern that heat injury might have occurred. Signs of heat injury or hyperalgesia, shown by profound erythema with pain or hypersensitivity of the skin, were identified as criteria for withdrawal from the study. No participants suffered any form of heat injury during this study.

Analysis

The receiver operating characteristic (ROC) curves of each stimuli pair and task were also plotted for every participant. The Gaussian unequal-variance model was fitted to the data using the RScorePlus software written by Lewis Harvey (http://psych.colorado.edu/~lharvey/html/software.html). RScorePlus is derived from Dorfman & Alf's (1969) RScore program and it provides a maximum-likelihood fit of the signal detection model to the rating data. There were altogether 36 ROCs (6 participants × 3 stimuli pairs × 2 tasks) generated for analysis. For the MDT, the adjacent temperatures were paired for analysis. This yielded the same number of stimulus pairs to the IRT. Data from both tasks were analyzed in a similar manner. The detection theory index of discriminability, d_a (Simpson & Fitter, 1973; Macmillan & Creelman, 2005) and the slopes of the ROCs based on three stimulus pairs, s were computed. The index d_a assumes an unequal variance model and is numerically equal to d' in the equal variance case. The Gaussian equal-variance index, d' was to be adopted if s for the discrimination and scaling data did not systematically depart from unity. When extreme response frequencies were present (i.e. categories containing proportions of zero), the categories were collapsed for analysis.

Cumulative sensitivity function

The d' values of adjacent stimuli for both tasks were cumulated to visualize the total sensitivity across the temperature range. Durlach & Braida (1969) named the resultant plots cumulative sensitivity functions (CSF). The lines of best fit through the origin were plotted using the least-squares method for the CSF of both tasks. The Weber fraction was calculated for each using the CSF. The Weber fraction, in this context, may be defined as the stimulus difference that is needed to produce a performance of $d' = 1$ as the just noticeable difference.

Relative judgmental variance

Equation 1 was used to estimate the relative variance, which is reproduced here as $\sigma^2_{MDT} / \sigma^2_{IRT} = \left(d'_{IRT} / d'_{MDT}\right)^2 - 1$, where d'_{MDT} is the sensitivity between the two adjacent temperature in the MDT, d'_{IRT} is the sensitivity between the two temperatures in the IRT, σ^2_{MDT} is the judgmental variance associated with the MDT, and σ^2_{IRT} is the stimulus variance associated with the IRT.

Results

Receiver Operating Characteristics

A total of 36 ROC curves were obtained. Out of these ROC curves, two curves significantly differed from the unequal variance model at the .05 significance level based on the chi-square goodness-of-fit statistic. The individual data from all subjects were jackknifed, based on the approach by Dorfman & Berbaum (1986), to generate 6 additional ROC curves to summarize the results of all stimulus pairs in both tasks. These

ROCs are shown in Figure 2. The jackknife procedure aims to avoid the common drawbacks of conventional averaging of sensitivity estimates (Macmillan & Kaplan, 1985). One of the drawbacks is obtaining a lower estimate of sensitivity compared to the original data if no averaging was used.

(Figure 2 about here)

The ROC slopes for the discrimination and scaling methods based on the three stimulus pairs were 1.01 (S.E. = .09) and 1.05 (S.E. = .13) respectively. The slopes for both tasks did not depart systematically from unity, therefore the Gaussian equal-variance d' was used instead of d_a.

Discriminability results

Figure 3 summarizes the discriminability of the stimulus pairs within each task. Although the data were jackknifed to generate the ROC curves in Figure 2, the conventional averaging of the sensitivity means was retained in Figure 3 to show the actual data for the six participants. Figure 3 showed that the average discriminability of the IRT task was always higher than the MDT task. This observation is as predicted by the analytical framework. It also agrees with results from previous results using noxious electrocutaneous stimulus (Irwin & Whitehead, 1991; Irwin et al., 1994; Rollman, 1983). Also, the discriminability of both tasks increased with an elevation of the temperatures of the stimulus pair. A repeated measures ANOVA (2 tasks x 3 stimulus pairs) performed on the discrimination ability data showed a significant main effect of task ($F(1,5) = 24.98$, p

= .004). There was also a significant main effect of stimulus pairs (*F*(2,10) = 5.37, *p* = .026,). Contrasts showed that sensitivity estimates for the 46-47°C stimulus pairs were not significantly higher than the 45-46°C stimulus pairs with a large effect size (*F*(1,5) = 2.63, *p* = .166, *r* = .59). The contrast also showed that sensitivity estimates for the 47-48°C stimulus pairs were significantly higher than the 45-46°C stimulus pairs with a large effect size (*F*(1,5) = 7.529, *p* = .041, *r* = .60). However, the interaction effect between task and stimulus pair was not significant (*p* = .152).

(Figure 3 about here)

Cumulative sensitivity functions

The cumulative sensitive functions (CSFs) were obtained using the jackknifed sensitivity estimates. The *d′* values of adjacent stimuli were cumulated. The successive cumulative sensitivities provided coordinates on the y-axis for plotting the CSF. Figure 4 shows the CSFs for this study. The linear functions were fitted to the data using least squares method with the functions passing through the origin. There is a difference between the slopes of the two CSFs. The slope for the IRT is steeper than the MDT, indicating that the overall discriminability of the IRT was better than the MDT. Since the linear fit of these functions were adequate, it may be said that the averaged discrimination performances of the participants were in accordance with Weber's law. The Weber fractions were found to be 0.026 for the MDT task and 0.015 for the IRT task.

(Figure 4 about here)

Relative judgmental variance

Based on Equation 1, the additional variance in the MDT was 2.18 times more than the IRT. This number was calculated using the cumulated sensitivity values obtained from the CSFs for the MDT and IRT.

Discussion

This study found that the MDT yielded decreased sensitivities compared to the IRT for noxious thermal stimuli. The amount of additional judgmental variance in the MDT was 2.18 times more than the IRT. These results were consistent with Durlach & Braida’s predictions (1969) and previous studies (Irwin & Whitehead, 1991; Irwin et al., 1994, Rollman, 1983).

The contribution of judgmental variance to a poorer sensitivity in MDT

This study was conducted to relate the direct scaling and discrimination methods (MDT and IRT respectively in this study) under a common framework. The present finding of lower sensitivity estimates yielded by the MDT as compared to the IRT supported the prediction that an additional component of variance may be attributed for direct scaling methods. Our results suggested that outcomes yielded from discrimination methods and direct scaling methods may be related. Therefore, this finding adds evidence to the assertion that discrimination methods were suitable for measuring responses from noxious stimuli. However, our results would have to be interpreted under the framework and assumptions of Durlach & Braida's (1969) theory. Irwin et al. (1994) stated that if this same analytical framework was extended to the method of magnitude estimation, similar results could be expected. It could be argued that the MDT used in this study is an example of scales that involves judgments of sensation magnitude (Braida & Durlach, 1972).

Lower relative variance in this study compared to previous studies

It is interesting to note that the relative variance between the two methods found in the present study was 2.18, lower than those found by Irwin & Whitehead (1991). The relative variance in their description (similar to the MDT) and identification task were 5.4 and 2.22 times more than the discrimination task respectively. This result is, perhaps, not unexpected, with two possible explanations. The first reason may be associated with the use of relatively lower trial numbers in this study, and the second reason may be connected with the stimulus range used for the MDT.

Influence of lower trial numbers on judgment variance

The introduction of lower trial numbers would inevitably increase both the variability of the responses and the likelihood of extreme proportions. This response variability may contribute considerable statistical bias to the sensitivity estimates (Hautus, 1997). An unpublished study by our research group found that when the number of trials per intensity in an one-interval, confidence-rating task was decreased from 40 to 17 trials, there was a 1.74 times increase in the amount of variance for the sensitivity estimates of the task using the lower trial number.

One might argue that higher trial numbers could be used to suppress the amount of variance in the sensitivity estimates and we acknowledge that this approach should be followed as closely as is practically possible. There are, however, also other considerations when large trial numbers are applied, such as the onset of heat injury and hyperalgesia (Pedersen & Kehlet, 1998), the ethical acceptability of prolonged noxious stimulation (International Association for the Study of Pain, 1995), and ultimately the transferability of the laboratory protocol to clinical studies. All of these factors should be carefully considered when deciding on the trial presentation numbers.

Influence of stimulus range on judgment variance

Another factor that may have influenced the amount of the judgment variance for the MDT was the range of stimuli judged. For the IRT, the participant was only required to concentrate on the difference between the stimuli pair presented. This is in contrast to the MDT, where participants may also have compared the sensation magnitude of the presented stimulus to the context of the stimulus range. This is despite the introduction of the trial-by-trial feedback provided to participants. A similar explanation was also offered by Rollman (1979) based on adaptation-level theory (Helson, 1964). Durlach & Braida (1969) theorized that when the stimulus range was large for the scaling task, the task became more difficult for the participants. has also This would lead to lowered sensitivity estimates. Since discrimination tasks, in Durlach & Braida's theory, are easier in terms of the absence of the judgmental component, the performance will always be better than the direct scaling tasks. However, they predicted that for a small stimulus range, the contribution of the judgment variance in direct scaling becomes almost negligible and performance on the scaling task would be similar to that on the discrimination task. This prediction was generally supported by Pynn, Braida & Durlach's (1972) study on auditory intensity discrimination. This raises another possible reason for the lower additional variance observed for the MDT task in our investigation compared to other studies. There is the possibility that the temperature range for this study was relatively narrow, thus causing smaller values of the judgment variance to be found, as predicted in Durlach & Braida's theory. Nevertheless, further studies need to be conducted to confirm this conjecture.

CSF as a potential tool for investigating suprathreshold sensitivities

The perception of noxious experimental stimuli has also been studied with methods obtaining point estimates of the transition between innocuousness and painfulness. (Graven-Nielsen, Sergerdahl, Svensson, & Arendt-Nielsen, 2001). An example of point estimates used in pain research is the determination of the pain threshold using the method of limits. The effectiveness of pain relief treatments have been largely evaluated based on the lowering of this threshold. It does not illuminate the effects of pain relief treatments on the suprathreshold sensitivities in which pain, the construct of interest, resides. This is especially important for suprathreshold sensitivities in studies examining nociception. The same criticism could be leveled at the sensitivity estimates obtained for individual stimulus pairs in this present study. The sensitivity estimates provided information confined only to that specific stimulus pair. However, it reveals little about the sensitivities contained within the sensory range of interest. This problem was solved, for the purposes of this study, through the use of cumulative sensitivity functions. Cumulative sensitivity functions may provide additional information on the suprathreshold range of sensitivities and the effects of intervention on them (Gracely, 2005). It may be a valuable tool for future studies investigating the description and influence of interventions on suprathreshold sensitivities.

The relevance of findings for future clinical studies

Our findings showed that when humans judged the intensities of thermal stimuli, decisions were made within the context of the type of task and stimulus range. Since most experimental measures of pain in clinical studies have used variants of the direct scaling method, our findings suggest that the responses of clinical participants may contain a component of judgment variance. In order to prevent diminishing the participants' discrimination ability by the effect of stimuli range comparisons (Poulton, 1989), the responses of clinical participants may be examined using the discrimination method, within the framework proposed by Irwin & Whitehead (1991).

Nevertheless, clinical pain is a multidimensional experience involving affective, cognitive and sensory components (Melzack, 1999). Chronic pain experienced by patients may be associated with changes in their empirical pain thresholds or self-reported pain intensity (Kosek, Ekholm & Hansson, 1996). The multicollinearity of affective, cognitive and sensory responses of pain may interact to alter the amount of variance within the psychophysical responses. Studies on affect and signal detection theory measures of pain scaling have provided some evidence that the latter may be influenced by affective disorders (Dworkin, Clark & Lipsitz, 1995; Kemperman et al., 1997). These studies have used the direct scaling method within the framework of signal detection theory. Since the direct scaling method yields an additional variance on participant responses, it would be interesting to establish, in future studies, the interaction between affect and pain response when in the absence of additional variance (i.e. using the discrimination method). A clinical study is currently being conducted by our research group examining this question in people with chronic pain.

An alternative interpretation: The dimensional hypothesis

Although we have interpreted the findings based on a theory of judgment, it is possible that the results could be due to dimensional overlap between the responses of both tasks. This would mean that the judgment theory may have to be revised for nociception because the judgment theory assumes perceptual one-dimensionality (Durlach & Braida, 1969; Macmillan & Creelman, 2005, p.113-115). This alternative interpretation of dimensional overlap could be tested in several ways. The first method would be to utilize a multidimensional analytical approach. For example, Clark, Yang, Carroll, & Janal (1986) and Clark, Ferrer-Brechner, Janal, Carroll & Yang (1989) have analyzed the dimensions of both experimental and clinical pain using Individual Differences Scaling (INDSCAL) procedures. Another method would be to observe the directional shifts of sensitivity from both discrimination and direct scaling methods when an analgesic or anesthetic procedure has been performed (Rollman, 1983). If the anesthetic procedure leads to similar directional shifts in sensitivity for both tasks, this provides some evidence that responses from both tasks exist on similar dimensions. A disconfirmation test for the perceptual dimension similarity hypothesis may also be investigated on painful clinical conditions. That is, some characteristics of the painful condition may interact with the experimental stimulus, which then yields opposite shifts in sensitivity between discrimination and direct scaling methods. Even so, disconfirmation does not negate the potential usefulness of both tasks for diagnostic purposes. In fact the underlying basis for the opposite shifts in sensitivity, be it biological or cognitive in nature, could be elucidated and applied as a powerful clinical diagnostic tool for future treatment of painful conditions.

Conclusion

This study demonstrated that the discrimination approach is comparable to the direct scaling approach. The bridge between the two approaches was possible through analyzing the data under the theoretical framework of Durlach & Braida (1969), based on the assumption of perceptual one-dimensionality. Our results were consistent with Durlach & Braida's prediction that an additional component of judgment variance contributed to the decreased sensitivity in the direct scaling approach. The finding is useful for clinical pain studies employing psychophysical methods of testing as well as informing diagnostic procedures for painful conditions. Regardless of the type of psychophysical method used in the clinical studies, it is possible to relate and compare the findings. This would also suggest that discrimination methods are admissible as psychophysical procedures for pain studies. Therefore, this framework may serve as a potentially useful tool for evaluating the often complex processes of pain perception.

References

BRAIDA, L.D., & DURLACH, N.I. (1972). Intensity perception. II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America,* 51*,* 483-502.

CATERINA, M.J., SCHUMACHER, M.A., TOMINAGA, M., ROSEN, T.A., LEVINE, J.D., & JULIUS, D. (1997). The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature,* 389*,* 816-824.

CLARK, W.C., FERRER-BRECHNER, T., JANAL, M.N., CARROLL, J.D., & YANG, J.C. (1989). The dimensions of pain: A multidimensional scaling comparison of cancer patients and healthy volunteers. *Pain,* 37*,* 23-32.

CLARK, W.C. YANG, J.C., CARROLL, J.D., & JANAL, M.N. (1986). Multidimensional scaling reveals two dimensions of thermal pain. *Journal of Experimental Psychology,* 12*,* 103-107.

CRAIG, J.C., & ROLLMAN, G.B. (1999). Somesthesis. *Annual Review of Psychology,* 50*,* 305-331.

DORFMAN, D.D., & ALF, E. Jr. (1969). Maximum-liklelihood estimation of parameters of signal-detection theory and determination of confidence intervals: Rating-method data. *Journal of Mathematical Psychology,* 6*,* 487-496.

DORFMAN, D.D., & BERBAUM, K. (1986). RSCORE-J: Pooled rating-method data: A computer program for analyzing pooled ROC curves. *Behavior Research Methods, Instruments & Computers,* 18*,* 452-462.

DURLACH, N.I., & BRAIDA, L.D. (1969). Intensity Perception. I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America,* 42*,* 372-383.

DWORKIN, R.H., CLARK, W.C., & LIPSITZ, J.D. (1995). Pain responsivity in major depression and bipolar disorder. *Psychiatric Research,* 56*,* 181.

GRACELY, R.H. (2005). Studies of pain in human subjects. In S.B. McMahon. & M. Koltzenburg (Eds.), *Wall and Melzack's Textbook of Pain* (5th ed., pp. 267-289). Philadelphia: Churchill Livingstone.

GRAVEN-NIELSEN, T., SERGERDAHL, M., SVENSSON, P., & ARENDT-NIELSEN, L. (2001). Methods for induction and assessment of pain in humans with clinical and pharmacological examples. In *Methods in pain research* (pp. 263-290). Boca Raton: CRC Press.

HAUTUS, M. J. (1997). Calculating estimates of sensitivity from group data: Pooled versus averaged estimators. *Behavior Research Methods, Instruments & Computers,* 29*,* 556-562.

HELSON, H. (1964). *Adaptation-Level Theory*. New York: Harper & Row.

INTERNATIONAL ASSOCIATION FOR THE STUDY OF PAIN (1995). Ethical guidelines for pain research in humans. *Pain,* 63*,* 277-278.

IRWIN, R.J., HAUTUS, M.J., DAWSON, N.J., WELCH, D., & BAYLY, M.F. (1994). Discriminability of electrocutaneous stimuli after topical anesthesia: Detection-theory measurement of sensitivity to painful stimuli. *Perception & Psychophysics,* 55*,* 132-152.

IRWIN, R.J., & WHITEHEAD, P.R. (1991). Towards an objective psychophysics of pain. *Psychological Science,* 2*,* 230-235.

KEMPERMAN, I., RUSS, M.J., CLARK, W.C., KAKUMA, T., ZANINE, E., & HARRISON, K. (1997). Pain assessment in self-injurious patients with borderline personality disorder using signal detection theory. *Psychiatric Research,* 70*,* 175-183.

KOSEK, E., EKHOLM, J., & HANSSON, P. (1996). Sensory dysfunction in fibromyalgia patients with implications for pathogenic mechanisms. *Pain,* 68*,* 375-383.

LAMING, D. (1984). The relativity of 'absolute' judgements. *British Journal of Mathematical & Statistical Psychology,* 37*,* 152-183.

LAMING, D. (1997). The psychophysical primitive. In *The measurement of sensation* (pp. 143-167). Oxford: Oxford University Press.

MACMILLAN, N.A., & CREELMAN, C.D. (2005). *Detection theory: A user's guide*. (2nd ed.) New Jersey: Lawrence Erlbaum Associates.

MACMILLAN, N.A., & KAPLAN, H.L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin,* 98*,* 185-199.

MELZACK, R. (1999). From the gate to the neuromatrix. *Pain,* Suppl. 6, S121-S147.

MEYER, R.A., & CAMPBELL, J.N. (1981). Evidence for two distinct classes of unmyelinated nociceptive afferents in monkey. *Brain Research,* 224, 149-152.

NAHRA H., & PLAGHKI L. (2005). Innocuous skin cooling modulates perception and neurophysiological correlates of brief CO_2 laser stimuli in humans. *European Journal of Pain,* 9, 521-530.

PEDERSEN, J.L., & KEHLET, H. (1998). Secondary hyperalgesia to heat stimuli after burn injury in man. *Pain,* 76, 377-384.

PERTOVAARA, A., MARTIKAINEN, I.K., HAGELBERG, N., MANSIKKA, H., NAGREN, K., HIETALA, J., & SHEININ, H. (2004). Striatal dopamine D2/D3 receptor availability with individual response characteristics to pain. *European Journal of Neuroscience,* 20, 1587-1592.

PETERSEN, K.L., & ROWBOTHAM, M.C. (2006). Quantitative testing scaled up for multicenter clinical research networks: A promising start. *Pain,* 123, 219-220.

POULTON, E.C. (1989). *Bias in quantifying judgments*. Hove: Lawrence Erlbaum.

PRICE, D.D., HU, J.W., DUBNER, R., & GRACELY, R.H. (1977). Peripheral suppression of first pain and central summation of second pain evoked by noxious heat pulses. *Pain,* 3, 57-68.

PRICE, D.D. (1994). Psychophysical measurement of normal and abnormal pain processing. In J. Boivie, P. Hansson, & U. Lindblom (Eds.), *Touch, temperature,* *and pain in health and disease: mechanisms and assessments* (pp. 3-25). Seattle: IASP Press.

PYNN, C.T., BRAIDA, L.D., & DURLACH, N.I. (1972). Intensity perception. III. Resolution in small-range identification. *Journal of the Acoustical Society of America,* 51*,* 559-566.

ROLLMAN G.B. (1977). Signal detection theory methods in pain research. *Pain,* 3*,* 187-211.

ROLLMAN, G.B. (1979). Signal detection theory pain measures: empirical validation studies and adaptation-level effects. *Pain,* 6*,* 9-21.

ROLLMAN, G.B. (1983). Multiple representations of experimental pain. In J.J. Bonica, U. Lindblom, & A. Iggo (Eds.), *Advances in pain research and therapy* (pp. 865-869). New York: Raven Press.

SIMPSON, A.J. & FITTER, M.J. (1973). What is the best index of detectability? *Psychological Bulletin,* 80*,* 481-488.

STAUD, R., PRICE, D.D., ROBINSON, M.E., MAUDERLI, A.P., & VIERCK, C.J. (2004). Maintenance of windup of second pain requires less frequent stimulation in fibromyalgia patients compared to normal controls. *Pain,* 110*,* 689-696.

TREEDE, R. D., MEYER, R. A., RAJA, S. N., & CAMPBELL, J. N. (1995). Evidence for two different heat transduction mechanisms in nociceptive primary afferents innervating monkey skin. *The Journal Of Physiology,* 483 (Pt 3)*,* 747-758.

BLANK

Figure Captions

Figure 1. A. Magnitude description scale representing the scaling method. This scale was presented to the participant during the magnitude description task for judgment. The participant verbally provided the number describing the magnitude of sensation felt. B. Intensity resolution scale representing the discrimination method. This scale was presented to the participant during the intensity resolution task for judgment. The participant verbally provided the number describing their degree of confidence on whether the stronger stimulus was presented.

Figure 2. Receiver operating characteristic (ROC) curves fitting using a jackknifed procedure utilizing the pooled ratings of all 6 subjects. Each panel shows the ROC curves of the MDT and IRT for each stimulus pair.

Figure 3. Mean discriminability, obtained through conventional averaging, of the MDT and IRT methods for all stimulus pairs. The error bars depict standard errors of the means.

Figure 4. Cumulative sensitivity functions for the MDT and IRT as a function of relative temperature. The jackknifed d' values were used to obtain the cumulative sensitivity functions.

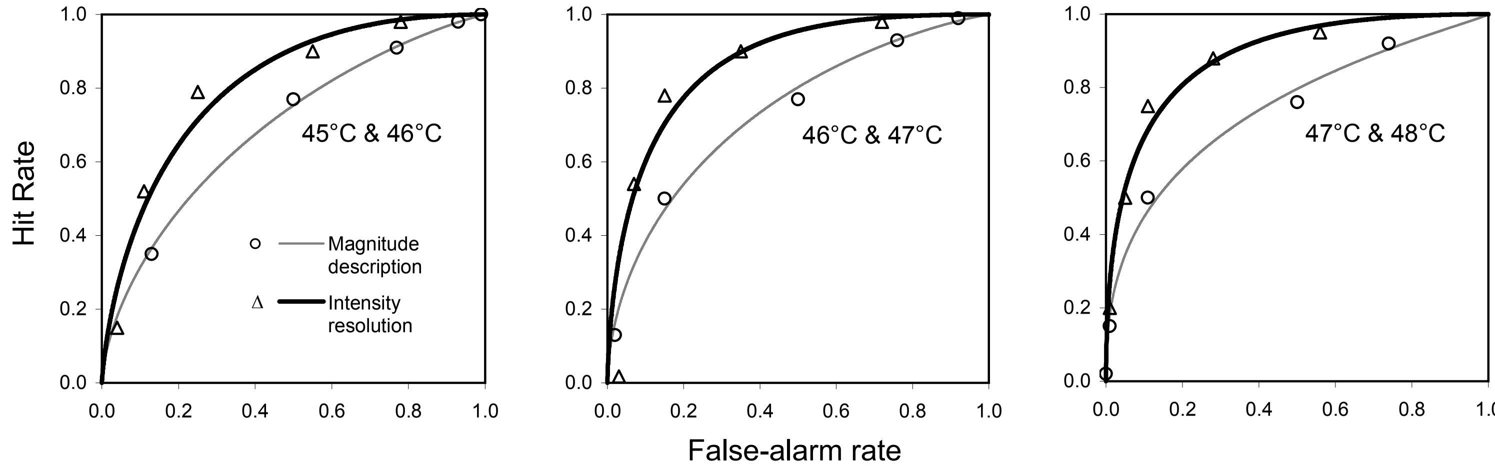
Figures

A

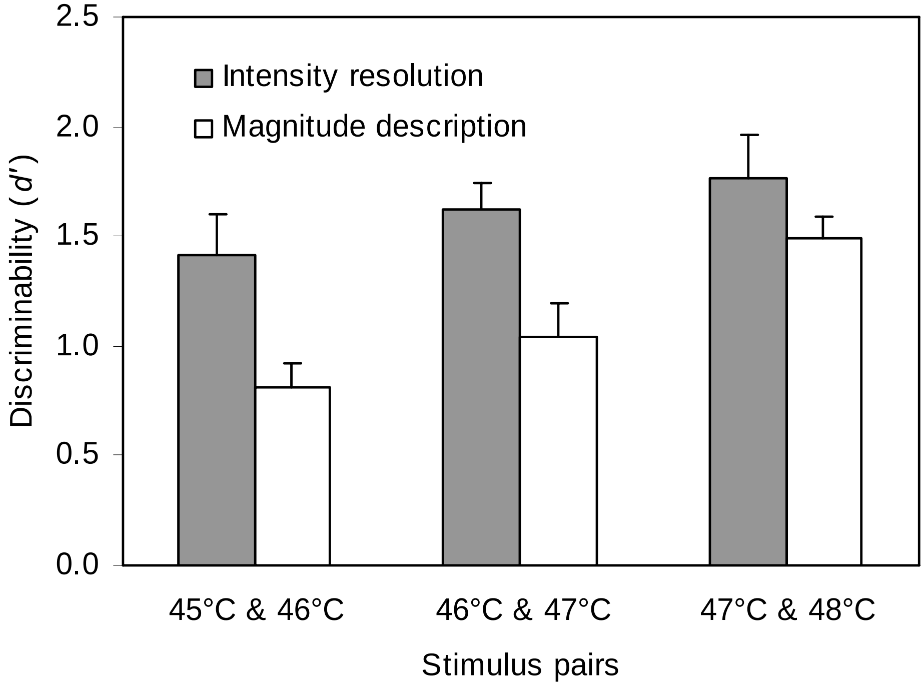
1	2	3	4	5	6
Warm	**Hot**	**Faint Pain**	**Painful**	**Very Painful**	**Severe Pain**

B

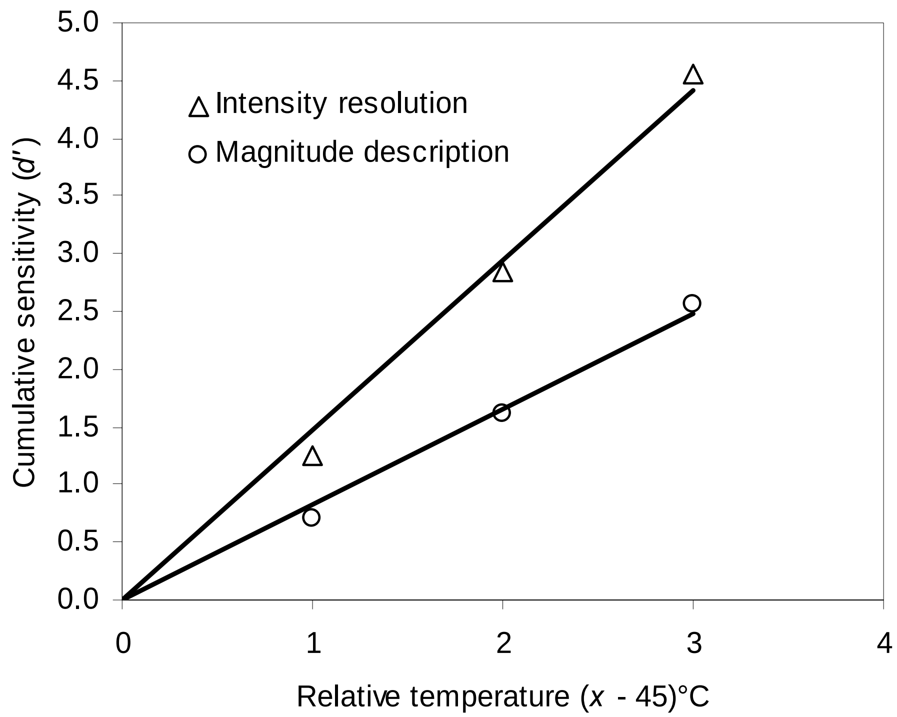
1	2	3	4	5	6
I am absolutely certain the weak stimulus was presented	**I am fairly certain the weak stimulus was presented**	**I am somewhat certain that the weak stimulus was presented**	**I am somewhat certain that the strong stimulus was presented**	**I am fairly certain the strong stimulus was presented**	**I am absolutely certain the strong stimulus was presented**



Hit Rate
False-alarm rate
45°C & 46°C
46°C & 47°C
47°C & 48°C
Magnitude description
Intensity resolution
0.0
0.2
0.4
0.6
0.8
1.0



2.5
2.0
1.5
1.0
0.5
0.0
Intensity resolution
Magnitude description
Discriminability (d')
45°C & 46°C
46°C & 47°C
47°C & 48°C
Stimulus pairs


Cumulative sensitivity (d')
Relative temperature (x - 45)°C
Intensity resolution
Magnitude description
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4