ORIGINAL ARTICLE
See related Commentary on pages v and xiv

Perceptual Matching for Assessment of Itch; Reliability and Responsiveness Analyzed by a Rank-Invariant Statistical Method

Elisabet Stener-Victorin,* Thomas Lundeberg,†‡ Jan Kowalski,† Lars Opdal,* Johan Sjöström,* and Lena Lundeberg§

*Institute of Occupational Therapy and Physiotherapy, The Sahlgrenska Academy, Göteborg University, Göteborg, Sweden; †Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden; ‡Rehabilitation Medicine, Karolinska Hospital, Stockholm, Sweden; §Department of Dermatology, Karolinska Hospital, Stockholm, Sweden

The aim of this study was to evaluate the reliability and responsiveness of a new device—perceptual matching by Matcher (Cefar Medical AB, Lund, Sweden)—in the assessment of the progress of experimentally induced itch as well as determine the reliability of the method in patients with skin disease and itch. The perceptual matching unit electrically stimulates the skin of the fingers in the left hand. When the amplitude of the sensation corresponded to that of the experienced itch, the subject was instructed to halt the stimulation and a value was automatically saved in the electronic device. A total of 36 healthy subjects and nine patients participated in the study. The healthy subjects were asked to rate the level of itch every 30 s during the first 5 min and thereafter every minute. The reliability was determined in a test–retest procedure of the time points 5, 10, and 15 min after induction of itching. To test the stability of the method, the first sensation of pain in healthy subjects without itch was determined every 10th minute during 60 min. In patients, the test–retest procedure was repeated within 5 min. Perceptual matching was found to be a reliable method of itch assessment, with no evidence for random individual disagreement between the assessments. The augmented rank order coefficient r_a was excellent: 1.00 at 5 min; 0.99 at 10 min; and 1.00 at 15 min. There was a clear indication of responsiveness for detecting changes in itch over time, $p \leqslant 0.05$. The perceptual matcher device can detect immediate changes in perceived itch and is also sensitive to gradual decreases after the induction of itch. The first sensation of pain in healthy subjects without itch was unaffected when assessed every 10th minute during 60 min. In patients, no evidence for random disagreement or systematic disagreement by group between repeated assessments was detected. The augmented rank order coefficient was high (0.98). In conclusion, Matcher measures itch intensity via perceptual matching with a high reliability and excellent responsiveness. The method is stable and can be recommended as an assessment tool for itch intensity both in experimental conditions as well as in patients with skin disease and itch. ***Key words: eczema/electrical stimulation/ histamine / itch / perceptual matching / reliability / responsiveness. J Invest Dermatol 121:1301–1305, 2003***

Itch (pruritus) is defined as an unpleasant sensory experience associated with the urge to scratch. Experimentally induced itch by iontophoresis is a reliable method and has been used widely to study the early stage of the cutaneous inflammatory responses flare and local edema (weal) (Magerl *et al*, 1990). Recently, microneurography has been used to describe subclasses of primary afferent C fibers that subserve distinct sensory functions, such as itch and touch (Schmelz *et al*, 1997; Vallbo *et al*, 1999; Olausson *et al*, 2002). One subset of these C-units is almost exclusively histamine sensitive and is probably the primary afferent whose activity elicits the itch sensation. This subset has been suggested to represent a new class of afferent nerve fibers (Schmelz *et al*, 1997). In addition, itch-specific central afferent neurons have recently been described in the dorsal horn (Andrew and Craig, 2001). It is reasonable to assume that itch and pain, physiologically, are transmitted in separate C and A-δ afferents. As an example, painful stimuli, e.g., electrical stimulation, can abolish histamine-induced itch locally for hours (Schmelz, 2001). Furthermore, morphine inhibits pain but can cause itch in humans and scratching behavior in animals (Thomas *et al*, 1993).

Reliability in the measurement of itch is of great importance, but the knowledge of the appropriate procedure is still scarce (Gracely, 1999a; Price, 2000). Similar to pain, itch is not a one-dimensional sensation and can be characterized by quantitative (i.e., itch intensity), qualitative (i.e., unpleasantness), and motor aspects (i.e., withdrawal, urge to scratch) (Drzezga *et al*, 2001). The visual analog scale (VAS) and different verbal and numeric scales are tools commonly used to assess itch intensity (Wahlgren *et al*, 1989). The VAS used for itch ratings is often carried out by an electronic device and stored in a computer for further analysis (Magerl *et al*, 1990; Schmelz *et al*, 1997; Koppert *et al*, 1998; Baron *et al*, 2001); however, some criticism of the use of the VAS has been voiced (Svensson, 2000; Lundeberg *et al*, 2001; Stener-Victorin *et al*, 2002). The VAS is bounded by fixed end-points and limits the range of measurements. The VAS is also a limited method of assessing the intensity as the patient must compare and grade perceived itch intensity against his or her worst-ever

Manuscript received March 31, 2003; revised July 2, 2003; accepted for publication July 24, 2003

Address correspondence and reprint requests to: Elisabet Stener-Victorin, PhD, RPT, Institute of Occupational Therapy and Physiotherapy, The Sahlgrenska Academy, Göteborg University, Box 455, SE-405 30 Göteborg, Sweden. Email: elsv@fhs.gu.se

experience of itch (Gracely, 1999b). VAS assessments offer a false impression of reliable measurements because the assessments are expressed in millimeters, but inconsistency and great random disagreement have been demonstrated (Svensson, 2000; Lundeberg *et al*, 2001).

Other techniques used to evoke a correlate for sensation and the perception of experimentally induced itch are based on psychophysical concepts such as touch, heat, or pinch stimulus (Handwerker and Kobal, 1993; Gracely, 1999b). Electrical stimulus against which the person may match the perceived itch with nonvisualized, predetermined end-points is another method for assessing itch (Handwerker and Kobal, 1993; Koppert *et al*, 1998; Gracely, 1999b). A recently developed perceptual matching device was designed for both clinical and experimental assessment of pain and is based on noninvasive electrical stimulation applied to the skin of one hand as a matching stimulus (Lundeberg *et al*, 2001; Stener-Victorin *et al*, 2002). This allows a continuous individual response with nonvisualized, predetermined lower or upper limits. Recent studies have found that the reliability of perceptual matching of pain experiences during a minor surgical operation is higher than that of the VAS (Stener-Victorin *et al*, 2002), whereas the responsiveness of the two methods is similar (Lundeberg *et al*, 2001; Stener-Victorin *et al*, 2002). The experiences of itch and pain have many similarities as both involve a complex response composed of sensory, affective, and cognitive aspects (Gracely, 1999a; Price, 2000). Perceptual matching and the present device have not previously been tested in the assessment of itch.

The main aim of this study was to evaluate the reliability (in a test–retest procedure) and the responsiveness (changes over time) of a new method—perceptual matching by Matcher, Cefar Medical AB in Lund, Sweden—for assessment of the progress of experimentally induced itch by using a rank-invariant statistical method. To test the stability of the method, the first sensation of pain in healthy subjects was assessed repeatedly during 60 min. In addition, reliability, i.e., test–retest, for the assessment of experienced itch was measured in patients with skin disease.

MATERIALS AND METHODS

Subjects and experimental design Thirty-six healthy subjects (18 women aged 20–48 y and 18 men, aged 21–42 y) were recruited among students at the Sahlgrenska Academy at Göteborg University, Sweden, to participate in this experimental study. None had any skin disease, histamine-induced allergy, asthma, diabetes, or pain. None consumed any medication that was intended for pain or that contained an antihistamine, and none were pregnant.

In addition, nine patients (four women aged 44–79 y and five men, aged 21–68 y) with a skin disease (eczema) and itch were recruited from the Department of Dermatology, Karolinska Hospital, Stockholm, Sweden. After the experimental procedure had been explained, all subjects gave written consent prior to the study. The local Ethics Committees of Göteborg and Stockholm Universities, Sweden, approved the study.

Experimental induction of itch Itch was induced by the iontophoresis of a 1% histamine diphosphate solution (Sigma-Aldrich Sweden AB, Stockholm, Sweden) in deionized water (Magerl *et al*, 1990). An oval pad (Iomed Trans QEsmall 1.5–2.0 mL) was filled with the histamine solution and connected to the anode of a commercial iontophoretic machine (PhoresorII Auto–IOMED, Salt Lake City, Utah). The electrode was applied 10 cm distal to the elbow fold on the subject's volar forearm, and a positive current of 2.0 mA was applied for 60 s via a silver–silver chloride electrode to deliver a standard amount of histamine. The current flowed through an indifferent electrode pad placed on the lateral side of the forearm. The iontophoresis procedure induced a sensation of itch in all subjects, lasting approximately 15 min.

Itch and first sensation of pain measured by perceptual matching Before histamine was applied, the subjects had no itch and the Matcher value was calibrated to 0. After the application of histamine (see details above), the subjects were asked to rate the level of perceived itch every 30 s for 5 min and thereafter every 60 s for a total of 20 min by electrically stimulating the skin, producing perceptual matching. At the time points 5, 10, and 15 min after application of histamine, the perceptual matching procedure was repeated within 30 s, test–retest, to test the reliability. When the amplitude of the sensation in the left hand corresponded to that of the experienced itch, the subject was told to halt the stimulation by pressing the button or releasing the fingers of the left hand from the electrode box. The value obtained was between 0 and 99 and automatically saved in the electronic device.

To test the stability of the perceptual matching procedure over time, the first sensation of pain was determined every 10th minute during 60 min in healthy subjects and they served as a negative control to experimental subjects. The subject was told to halt the stimulation, as described above, when the first sensation of pain was perceived in the left hand.

In addition, patients with skin disease and itch were asked to rate the level of perceived itch. The perceptual matching procedure was repeated 3 times test–retest with 5 min between the assessments, to test the reliability of the Matcher.

Electrical stimulation unit Matcher is an instrument for registering, assessing, and quantifying itch intensity using constant current electrical stimulation (CCES). The unit is programmed to produce a constant current despite variations in skin resistance (e.g., influenced by sweating and anxiety of the subject) up to 13 kΩ. The patient holds the instrument in a pincer grasp between the thumb and index finger on the left hand. The fingertips press against the electrodes on the instrument. The patient starts and interrupts stimulation by gripping or releasing the electrodes or by pressing the button. The generator detects the resistance and regulates the current accordingly, which means that the instrument always maintains an amplitude of 15 mA, regardless of the resistance. The stimulation, a monophasic square pulse with a frequency of 10 Hz, starts out at zero and then increases continuously by lengthening the pulse duration. The initial duration is 4 μs and the maximum 396 μs. The increase takes place in increments of 4 μs per step up to a total of 99 steps. Three different velocities between 0 and 250 μs with a duration of 35, 40, or 45 s are randomly selected. The value presented on the Matcher (0–99) is related to the pulse width by a data transformation by a division of 4 and is displayed on a liquid crystal display screen and saved in the memory. The electrical charge per second is extremely low and varies through the different steps from 0.6 to 59 μC per s.

Because of the electrical stimulation, persons with a pacemaker should not use this method.

Statistics All data were presented using descriptive statistics. They include numerical units assessed by the Matcher as median values and ranges from minimum to maximum. All statistical calculations were made using the software package SYSRAN 1.0 for Matlab 6 (JK Stat, Stockholm, Sweden).

The data were considered to have an ordered structure. No metric properties were assumed. The data were therefore characterized as ordinal data (Svensson, 2000). The statistical approach to this study considers the rank-invariant properties of ordinal data without any other structural assumptions. The advantage with the rank-invariant statistical method is that the results are unaffected by the data transformation performed on the original data from the Matcher. Aspects of both systematic and random disagreement could also considered.

To estimate the reliability (agreement) between two repeated observations (test–retest) in itch assessment by perceptual matching, the measures of systematic disagreement—relative position (RP) and relative concentration (RC)—were estimated together with the measure of random disagreement (RV). Systematic disagreement is caused by systematically increased or decreased values as a group from the first to the second assessment, i.e., most of the subjects rate their itch systematically higher or lower at the repeated assessment. The random individual disagreement is a measure of disagreement on the individual level that could not be explained by systematic disagreement. Values of RP and RC range from −1 to 1, where 0 indicates a lack of systematic disagreement between the two repeated assessments. RV values close to 0 indicate a lack of random disagreement. The higher the values of RV, the more dispersed are the observations (Svensson, 2000). An RV-related measure to express the size of random individual changes is the rank-order agreement coefficient, r_a, where values close to 1 indicate a lack of random disagreement.

Systematic disagreement was graphically illustrated by plotting the cumulative proportions of observed threshold values against each other starting at the point 0.0. This illustration defines a relative operative characteristic (ROC) curve. The ROC curve coincides with the main diagonal when there is a lack of bias, indicating no systematic disagreement.

The same statistical method was used to analyze the responsiveness to induction of itch. Responsiveness is a measure of how much of the

changes in itch measures could be identified as a systematic change by group. The changes not explained by responsiveness are defined as random individual changes. In this case, the measure of RP indicates a systematic change by group between 5 and 10 min and between 10 and 15 min occasions. Individual heterogeneity in changes are identified by RV. The approximate 95% confidence intervals (CI) for the estimates were calculated using the jack-knife estimation of standard errors. A p-value less than 0.05 was considered significant.

RESULTS

Reliability of itch assessments by perceptual matching of experimentally induced itch **Figure 1** shows the ROC curves for repeated (test–retest) perceptual matching measurements at 5, 10, and 15 min after induction of itch. The ROC curve more or less follows the main diagonal. This indicates a lack of systematic disagreement in repeated itch assessments. **Table I** shows the pattern of disagreement in repeated observations of perceived itch. No statistical evidence for systematic disagreement was found in the RP and RC values. The RV values indicate a lack of random individual disagreement in itch in all three assessments. The augmented rank order agreement coefficient, r_a was excellent: 1.00 at 5 min, 0.99 at 10 min, and 1.00 at 15 min, which indicates that the reliability between repeated assessments is high.

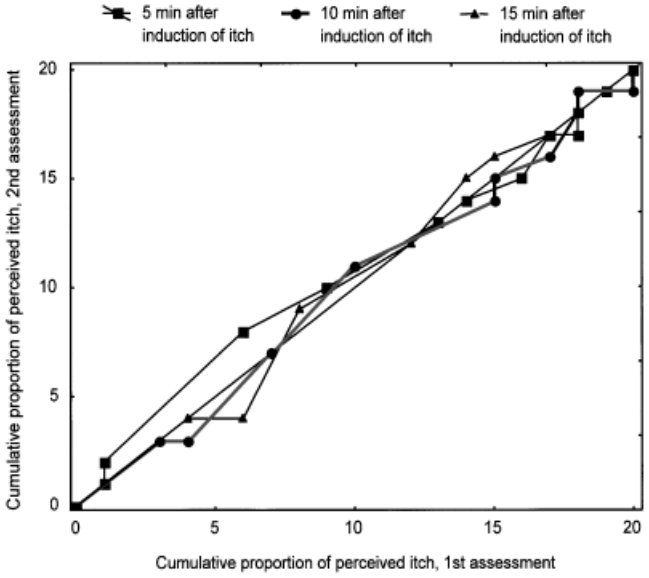

Figure 1. The relative operating characteristic (ROC) curves illustrate the systematic disagreement of repeated measurements between the first and second (test–retest) perceptual matching measurements recorded at 5, 10, and 15 min after induction of itch.

Responsiveness of itch assessments by perceptual matching of experimentally induced itch **Figure 2** shows the median value, interquartile range, and minimum and maximum values for repeated itch matching assessments by perceptual matching during the experimental procedure. The levels of perceived itch increased significantly immediately (1 min) after induction of itch ($p < 0.001$). **Table II** shows the pattern of changes in itch intensity for perceptual matching at 5, 10, and 15 min after the induction of itch. The RP estimates indicate a systematic decrease in itch intensity between the 10 and 15 min registrations, which was statistically confirmed, $p \leqslant 0.05$. Perceived itch intensity levels were reduced but present at the 20 min registrations.

First sensation of pain measured by perceptual matching **Figure 3** shows the median value, interquartile range, and minimum and maximum values for repeated assessments of the first sensation of pain in healthy subjects without itch by perceptual matching. There was no statistical evidence for systematic changes in the first sensation of pain in the repeated measures taken at 10 min intervals during 60 min.

Reliability of itch assessments by perceptual matching in patients with skin disease The results for systematic and individual disagreement for the three repeated assessments of itch by perceptual matching in patients with skin disease are

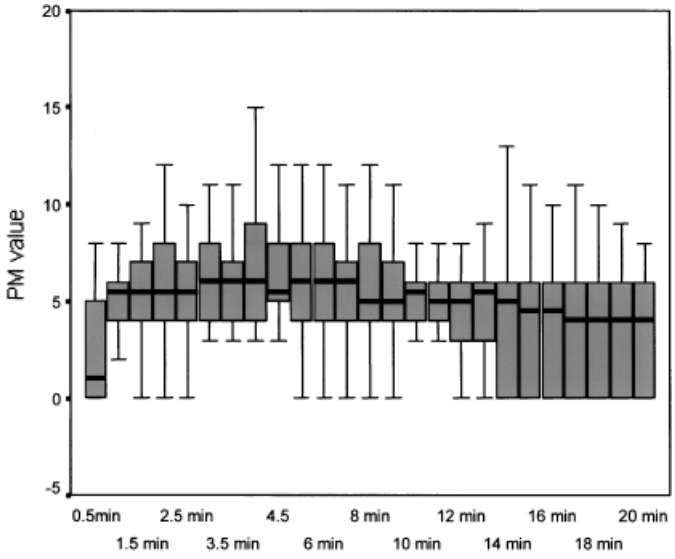

Figure 2. The median value, interquartile range, and minimum and maximum values for repeated itch matching assessments by perceptual matching during the experimental procedure.

Table I. The measures of group and individual disagreement in the assessment of repeated observations (test-retest) of perceived itch using perceptual matching 5 min, 10 min, and 15 min after induction of itch. The measures and standard error (SE) are presented together with their 95% confidence intervals (CI).

		Time after induction of itch (min)		
		5	10	15
Group systematic disagreement				
in position	RP (SE)	−0.06 (0.043)	0.01 (0.042)	−0.012 (0.043)
	CI %	−0.144, 0.024	−0.072, 0.092	−0.097, 0.073
in concentration	RC (SE)	−0.137 (0.067)	−0.008 (0.058)	0.078 (0.056)
	CI %	−0.268, −0.007	−0.122, 0.106	−0.031, 0.187
Random individual disagreement	RV (SE)	0 (0)	0.012 (0.015)	0.002 (0.003)
	CI %	0, 0	−0.018, 0.042	−0.004, 0.008
Augmented rank order agreement coefficient	r_a	1.00s	0.99	1.00

RP = Relative Position; RC = Relative Concentration; RV = Random disagreement

Table II. The measures of group and individual changes in the assessment of the development of experimentally induced itch with perceptual matching between 5 and 10 min and between 10 and 15 min. The measures and standard error (SE) are presented together with their 95% confidence intervals (CI).

		Measured itch	
		Itch 5 min vs 10 min	Itch 10 min vs 15 min
Systematic change for the group			
in position	RP (SE)	−0.06 (0.118)	−0.301 (0.118)
	CI %	−0.291, 0.171	−0.536, −0.075
in concentration	RC (SE)	0.048 (0.160)	−0.064 (0.146)
	CI %	−0.266, 0.362	−0.349, 0.222
Random individual changes	RV (SE)	0.256 (0.173)	0.222 (0.161)
	CI %	−0.833, 0.596	−0.093, 0.538
Augmented rank-order agreement coefficient	r_a	0.74	0.78

RP = Relative Position; RC = Relative Concentration; RV = Random disagreement

Table III. The measures of group and individual disagreement for the three repeated assessments of perceived itch using perceptual matching between 0 and 5 min and between 5 and 10 min in patients with skin disease. The measures and standard error (SE) are presented together with their 95% confidence intervals (CI).

		First and second assessment	Second and third assessment
Group systematic disagreement			
in position	RP (SE)	0.061 (0.072)	−0.012 (0.132)
	CI %	−0.080, 0.203	−0.270, 0.246
in concentration	RC (SE)	0.086 (0.147)	−0.183 (0.172)
	CI %	−0.203, 0.375	−0.520, 0.153
Random individual disagreement	RV (SE)	0.016 (0.027)	0.165 (0.208)
	CI %	−0.038, 0.070	−0.244, 0.573
Augmented rank-order agreement coefficient	r_a	0.98	0.84

RP = Relative Position; RC = Relative Concentration; RV = Random disagreement

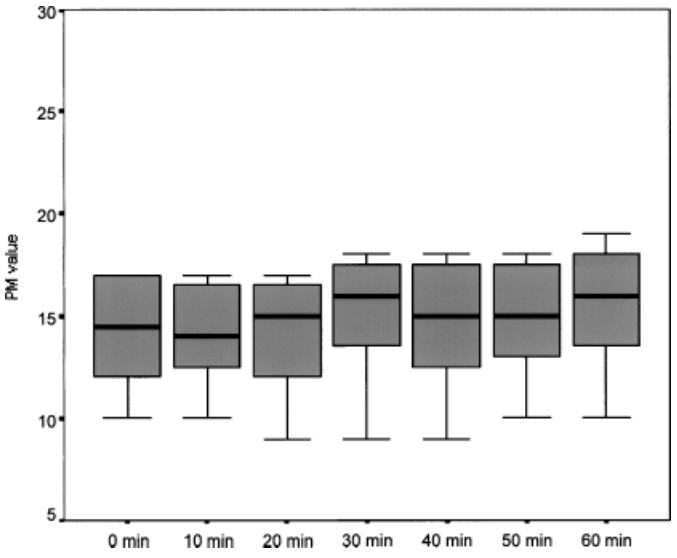

Figure 3. The median value, interquartile range, and minimum and maximum values for repeated assessments of the first sensation of pain by perceptual matching every 10th minute during 60 min.

presented in **Table III**. There was no evidence of random disagreement or systematic disagreement by group between repeated assessments. The augmented rank order coefficient r_a was excellent (between 0.98 and 0.84).

DISCUSSION

The main findings in this study are that perceptual matching by the Matcher measures itch with high reliability (agreement) in the test–retest situation in both experimentally induced itch as well as in patients with skin disease and itch. Furthermore, the perceptual matching method is sensitive and detects true changes in itch over time (responsiveness).

The reliability of perceptual matching by electrical stimulation was excellent in the test–retest evaluations at the 5, 10, and 15 min assessments of itch intensity, as shown by no systematic disagreement as a group in position (RP) or in concentration (RC). The velocity of the electrical pulses experienced by the subjects was random, and the values were unknown beforehand. Therefore, any expectations of the subject or the experimenter had no effect on the outcome. In addition, in patients with skin disease and itch, the reliability, i.e., test–retest assessments of itch by perceptual matching, was demonstrated as good in subjects with experimentally induced itch, which makes it attractive for clinical use.

Furthermore, the results of this study found no random individual changes (RV) in the estimates of responsiveness to experimental induction of itch when itch assessment was made using the method of perceptual matching and electrical stimulation described in this study. The low RV value found indicates that the potential for perceptual matching to detect true changes in perceived itch—responsiveness—is great. This was shown by a gradual increase followed by a decrease in perceptual matching levels.

In tests of the first sensation of pain every 10th minute during 60 min in healthy subjects without itch, no changes over time were observed, demonstrating that the perceptual matching method is stable and that there is no tendency for a training effect or sensitization to develop.

The VAS has traditionally been used as a method of assessing the progress of histamine-induced itch and evaluating different interventions. VAS is a limited and less powerful method in several ways, however, as previously discussed (Svensson, 2000; Lundeberg *et al*, 2001; Stener-Victorin *et al*, 2002), and its usefulness has been discussed (DeLoach *et al*, 1998). New methods for itch assessment are needed, and electrical stimulation, against which the subject may match the perceived itch, is a good alternative method (Lundeberg *et al*, 2001; Stener-Victorin *et al*, 2002).

The rank-invariant statistical method used in this study has greater advantages than more traditional nonparametric methods, e.g., kappa statistics, when evaluating agreement. The method divides the disagreement into systematically and randomly caused errors. This is important when attempting to identify the mechanism behind an unreliable method of measurement.

Perceptual matching was well accepted by both the study subjects and the patients and no complications were reported. Safety aspects are vital and must be considered for ethical reasons when a new method based on electrical stimulation is introduced, as it could be perceived as harmful to tissue. For that reason, it is important that the subject is able to terminate stimulation immediately at any time in the procedure. The sensation is brief and perceived by most subjects as pricking, throbbing, or tingling, although some may find it uncomfortable.

In conclusion, use of the Matcher to measure itch intensity by perceptual matching is highly reliable with excellent responsive-

ness. The method is stable and is recommended as an assessment tool for itch intensity, both in experimental conditions as well as in patients with skin disease and itch.

We thank the volunteers and patients who participated in this study. CefarMatcher AB (Lund, Sweden) supported this work by lending the perceptual matching equipment. The Wilhelm and Martina Lundgrens's Science Fund, Jubileumsfonden (Göteborg University), and the Hjalmar Svensson Foundation provided funding.

REFERENCES

Andrew D, Craig AD: Spinothalamic lamina I neurons selectively sensitive to histamine: A central neural pathway for itch. *Nat Neurosci* 4:72–77, 2001

Baron R, Schwarz K, Kleinert A, Schattschneider J, Wasner G: Histamine-induced itch converts into pain in neuropathic hyperalgesia. *Neuroreport* 12:3475–3478, 2001

DeLoach LJ, Higgins MS, Caplan AB, Stiff JL: The visual analog scale in the immediate postoperative period: Intrasubject variability and correlation with a numeric scale. *Anesth Analg* 86:102–106, 1998

Drzezga A, Darsow U, Treede RD, *et al*: Central activation by histamine-induced itch: Analogies to pain processing: A correlational analysis of O-15 H2O positron emission tomography studies. *Pain* 92:295–305, 2001

Gracely RH: Pain measurement. *Acta Anaesthesiol Scand* 43:897–908, 1999a

Gracely RH: Studies of pain in normal man. In: Wall PD, Melzack R (eds). *Textbook of Pain*. London: Churchill Livingston, 1999b; p 385–407

Handwerker HO, Kobal G: Psychophysiology of experimentally induced pain. *Physiol Rev* 73:639–671, 1993

Koppert W, Zeck S, Sittl R, Likar R, Knoll R, Schmelz M: Low-dose lidocaine suppresses experimentally induced hyperalgesia in humans. *Anesthesiology* 89:1345–1353, 1998

Lundeberg T, Lund I, Dahlin L, *et al*: Reliability, responsiveness of three different pain assessments. *J Rehab Med* 33:1–5, 2001

Magerl W, Westerman RA, Mohner B, Handwerker HO: Properties of transdermal histamine iontophoresis: Differential effects of season, gender, and body region. *J Invest Dermatol* 94:347–352, 1990

Olausson H, Lamarre Y, Backlund H, *et al*: Unmyelinated tactile afferents signal touch and project to insular cortex. *Nat Neurosci* 29:1–5, 2002

Price DD: Psychological and neural mechanisms of the affective dimension of pain. *Science* 288:1769–1772, 2000

Schmelz M: A neural pathway for itch. *Nat Neurosci* 4:9–10, 2001

Schmelz M, Schmidt R, Bickel A, Handwerker HO, Torebjork HE: Specific C-receptors for itch in human skin. *J Neurosci* 17:8003–8008, 1997

Stener-Victorin E, Kowalski J, Lundeberg T: A new highly reliable instrument for the assessment of pre- and postoperative gynecological pain. *Anesth Analg* 95:151–157, 2002

Svensson E.: Concordance between ratings using different scales for the same variable. *Stat Med* 19:3483–3496, 2000

Thomas DA, Williams GM, Iwata K, Kenshalo DR Jr, Dubner R: Multiple effects of morphine on facial scratching in monkeys. *Anesth Analg* 77:933–935, 1993

Vallbo AB, Olausson H, Wessberg J: Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. *J Neurophysiol* 81: 2753–2763, 1999

Wahlgren CF, Ekblom A, Hagermark O: Some aspects of the experimental induction and measurement of itch. *Acta Derm Venereol* 69:185–189, 1989