Cs

Evaluating the 'Labeled Magnitude Scale' for Measuring Sensations of Taste and Smell

Barry G. Green, Pamela Dalton, Beverly Cowart, Greg Shaffer, Krystyna Rankin and Jennifer Higgins

Monell Chemical Senses Center, Philadelphia, PA 19104–3308, USA

Correspondence to be sent to: Barry G. Green, John B. Pierce Laboratory, 290 Congress Avenue, New Haven, CT 06519, USA

Abstract

The Labeled Magnitude Scale (LMS) is a semantic scale of perceptual intensity characterized by a quasi-logarithmic spacing of its verbal labels. The LMS had previously been shown to yield psychophysical functions equivalent to magnitude estimation (ME) when gustatory, thermal and nociceptive stimuli were presented and rated together, and the upper bound of the LMS was defined as the 'strongest imaginable oral sensation'. The present study compared the LMS to ME within the more limited contexts of taste and smell. In Experiment 1, subjects used both methods to rate either taste intensity produced by sucrose and NaCl or odor intensity produced by acetic acid and phenyl ethyl alcohol, with the upper bound of the LMS defined as either the 'strongest imaginable taste' or the 'strongest imaginable odor'. The LMS produced psychophysical functions equivalent to those produced by ME. In Experiment 2 a new group of subjects used both methods to rate the intensity of three different taste qualities, with the upper bound of the LMS defined as the 'strongest imaginable [sweetness, saltiness, or bitterness]'. In all three cases the LMS produced steeper functions than did ME. Experiment 3 tested the hypothesis that the LMS yields data comparable to ME only when the perceptual domain under study includes painful sensations. This hypothesis was supported when the LMS again produced steeper functions than ME after subjects had been explicitly instructed to omit painful sensations (e.g. the 'burn' of hot peppers) from the concept of 'strongest imaginable taste'. We conclude that the LMS can be used to scale sensations of taste and smell when they are broadly defined, but that it should be modified for use in scaling specific taste (and probably odor) qualities. The implications of these results for theoretical issues related to ME, category-ratio scales and the size of the perceptual range in different sensory modalities are discussed. **Chem. Senses 21: 323–334, 1996.**

Introduction

The Labeled Magnitude Scale (LMS) is a semantically-labeled scale of sensation intensity that was developed for the study of oral somatosensation and gustation (Green *et al.*, 1993). The LMS is based upon the concept of the 'category ratio' (CR) scale, which Borg (1982) initially devised for the measurement of perceived exertion. Because its category labels were positioned according to their semantic magnitudes, Borg argued (Borg, 1982, 1990; Borg and Borg, 1987) that, in addition to providing information about absolute intensity, the CR scale could also yield ratio-level scaling data equivalent to that produced by magnitude estimation (ME) (Stevens and Galanter, 1957; Stevens, 1971). Borg further asserted, as Teghtsoonian (1971, 1973) had before him, that the perceptual range is similar in all sensory modalities and, hence, that the CR scale could be used to scale sensations other than exertion. To test this hypothesis, Borg and his colleagues compared psychophysical functions obtained with the CR scale for exertion, taste and loudness with functions obtained with ME. Agreement between methods was better for perceived exertion than for the other modalities, even after the CR scale had been adjusted in various ways (Borg, 1982, 1990; Marks *et al.*, 1983; Borg *et al.*, 1985; Borg and Borg, 1987).

The questionable performance of the CR scales in contexts other than perceived exertion prompted Green *et al.* (1993) to develop the LMS. The scale was constructed by asking 33 subjects to give magnitude estimates to verbal descriptors of intensity (e.g. 'weak', 'strong') that were interspersed among written examples of a variety of common oral sensations (e.g. 'the coldness of an ice cube', 'the saltiness of soup'), which the subjects also rated. The resulting scale, shown in Figure 1, is composed of seven verbal labels arranged according to the geometric means of their rated magnitudes. The key features of the LMS are the unequal, quasi-logarithmic spacing of its verbal labels and the presence of 'strongest imaginable' at its upper bound. A comparison of the LMS with the various versions of the CR scale indicated that the LMS encompassed a wider numerical range between its lowest ('barely detectable') and highest ('strongest imaginable') verbal descriptors. To determine whether the LMS could produce intensity data comparable to ME for a variety of oral stimuli, subjects who had no

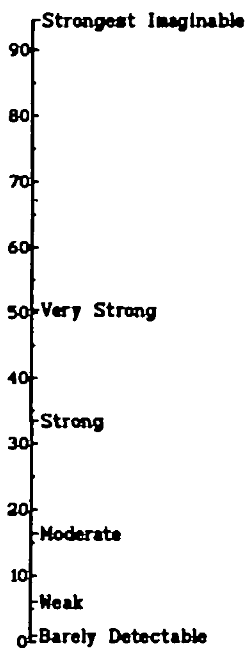

Figure 1 The labeled magnitude scale devised by Green *et al.* (1993) that was evaluated in the present study. The scale was presented on a computer screen with the numbers and tick marks removed.

experience with psychophysical scaling used the two methods to rate the intensities of sensations of taste (sweetness of sucrose), chemesthesis (sensory irritation from ethanol) and temperature (cold water). Ratings of intensity on the LMS were made relative to the 'strongest imaginable' oral sensation of any kind. After normalization to eliminate the effects of idiosyncratic number usage in ME, the psychophysical functions generated by the two methods were statistically indistinguishable (Green *et al.*, 1993), which implies that the LMS can yield ratio-level data (Stevens, 1971; Marks, 1974) on the intensity of diverse oral sensations when the sensations are experienced and judged within a common perceptual context.

Although the results from the first study were promising, fundamental questions remained regarding the generalizability of the LMS. Of particular interest was whether the LMS could be used as a specific 'taste scale' or 'smell scale'. That is, would the LMS produce psychophysical functions comparable to those obtained with ME if subjects judged tastes and odors in the context of the 'strongest imaginable taste' or the 'strongest imaginable odor' rather than in the context of the 'strongest imaginable oral sensation'? This question has both practical and theoretical significance. From a practical standpoint, as currently used, the LMS permits interpretation of perceived intensity only in terms of 'oral sensation'. This limitation prohibits conclusions about the intensity of gustatory or olfactory stimuli within the perceptual domains of taste and smell, e.g. whether a stimulus produces a 'weak' or 'strong' taste. In addition, because intense tastes and smells would rarely reach or exceed 'strong' on a scale of all imaginable oral sensations (which include intense pains), ratings of gustatory and olfactory stimuli would tend to be confined to the lower portion of the LMS. If the top of the scale were redefined as the 'strongest imaginable taste [or odor]', responses should be distributed over more of the LMS, which would presumably provide finer-grained information about the intensity of gustatory (or olfactory) sensations. From a theoretical standpoint, equivalence between the LMS and ME can occur only if the numerical ratio between 'barely detectable' and 'strongest imaginable' corresponds to the full range of conceivable sensations on the perceptual continuum of interest. Without such correspondence, ratios among sensations would tend to be overestimated if the response scale were too large and underestimated if the response scale were too small. Accordingly, the LMS offers another opportunity to evaluate the Teghtsoonian–Borg hypothesis that the perceptual range is invariant. Given that the response range of the LMS is fixed, failure to find agreement with ME in one or more sensory domains would argue against the invariance hypothesis. Although agreement between methods had been found when the intensity of three different kinds of sensations were rated within the broad context of all oral sensations (Green *et al.*, 1993), it remained to be learned whether agreement would be maintained if intensity was rated within the limited domains of taste and smell.

Accordingly, the present study was carried out to discover if redefining the upper bound of the LMS would enable it to be used within the narrower contexts of all tastes or all smells, or within the still narrower context of individual taste qualities. Three experiments were conducted that compared the LMS with ME: the first experiment showed that the LMS produced psychophysical functions comparable to ME when the intensities of odors or tastes were measured within the context of the 'strongest imaginable odor sensation' or the 'strongest imaginable taste sensation', but the second experiment yielded discordant results for the two methods when intensity was rated within the specific taste qualities of sweetness, saltiness and bitterness. The third experiment provided evidence that equivalence between the LMS and ME holds only when the subject's concept of the 'strongest imaginable sensation' on a continuum includes painful sensations.

Experiment 1: smell and taste

The purpose of this experiment was to determine if the LMS could be used specifically to measure the intensity of odors or tastes. To provide a sufficiently rigorous and generalizable test, two different stimuli were tested within each sensory modality: one believed to affect only the gustatory or olfactory system, and one believed to stimulate the trigeminal nerve as well. Because it was necessary to use inexperienced subjects to compare the results from the LMS and ME, scaling of odors and tastes was conducted independently in two groups of subjects.

Methods

Odor scaling

Subjects. Eighteen people (10 women and eight men, median age = 25 years), none of whom had experience with either scaling method, were paid to participate. Subjects were free of head colds or allergies, and were screened for the ability to smell the test odorants. Subjects were asked to identify the 'different' bottle among three when one bottle contained the middle concentration of the test series and the other two were blanks. Successful discriminations on each of two trials with each test odorant were required for inclusion in the experiment.

Stimuli. The stimuli tested were phenyl ethyl alcohol (PEA), which has a floral smell and in vapor phase is believed to stimulate only the olfactory system in humans (Doty *et al.*, 1978; Kobal and Hummel, 1988) and acetic acid (vinegar), which is a common nasal and oral irritant. Both stimuli were delivered in five concentrations plus one blank: 100, 10, 1, 0.1 and 0.01% phenyl ethyl alcohol dissolved in glycerol, and 0.078, 0.039, 0.019, 0.0097 and 0.0048% acetic acid dissolved in propylene glycol. The solvents served as the blanks.

The stimuli were delivered in 240-ml glass bottles equipped with flip-top caps that had been modified to accommodate both a conical Teflon nose piece and a Teflon air intake tube. To sample an odorant the subject placed the nose piece snugly against one naris, blocked the other naris with a finger and inhaled the headspace with a normal, 2–3 s sniff. The air intake tube, which extended 11.5 cm into the bottle, enabled fresh air to flow toward the bottom of the bottle as the headspace was drawn off through the nose piece.

Psychophysical procedure. Each subject was tested individually in two sessions conducted on separate days, a maximum of 3 days apart. The stimuli were rated in one session using the LMS and in the other session using ME, with half of the subjects tested first with the LMS and half tested first with ME. Both odorants were presented in both sessions in a mixed design with the order of presentation randomized across concentration. Each concentration was presented twice, and the inter-stimulus interval (ISI) was approximately 30 s.

Anterior rhinometry was used to measure the patency of each naris three times per session: before the first stimulus presentation, midway through the stimulus trials and after the final stimulus. If airflow differed between nostrils by more than 20%, the subject was instructed to present subsequent stimuli to the more patent side. Otherwise, the subject presented all stimuli to the preferred nostril.

The position of the verbal labels on the LMS, as percentages of full scale length, are: barely detectable, 1.4; weak, 6.1; moderate, 17.2; strong, 35.4; very strong, 53.3; strongest imaginable, 100. The instructions for use of the LMS were the same as had been used in the first study that compared the LMS to ME (Green *et al.*, 1993), except that the context within which ratings were to be made was limited to all possible odors rather than to all possible oral sensations. The following paragraph in the LMS instructions defined the context in which ratings were to be made:

> In making your judgements of the intensity of odor sensations, you should rate the stimuli relative to other odor sensations of all kinds that you have experienced. This includes such varied sensations as the strong smell of rotting garbage, the odor from a subtle perfume, or the tingling/burning from smoky air. Thus, 'strongest imaginable' refers to the most intense odor sensation that you can imagine experiencing.

The LMS was displayed vertically on a computer monitor and responses were made by moving a cursor to the appropriate location on the scale using a mouse. Subjects were invited to ask questions about how to use the scale, but no practice trials were given.

The instructions for ME emphasized that after assigning a number to the first sensation, subsequent ratings should be based on the relative intensities of the sensations. The specific instructions pertaining to ratio-estimation were the following:

> For example, the first stimulus you receive might produce a moderate sensation, to which you should assign some moderate number—any one that seems appropriate to you. You should keep that number and the sensation you assigned it to in mind when you rate the next sensation. So if you assigned, for instance, a 7, and the next sensation was twice as strong, you should call it 14. If it is only half as strong, you should call it 3.5. On subsequent trials assign numbers similarly, always comparing the present sensation to the ones before.

It was further emphasized that subjects should not think in terms of a fixed range of numbers, such as the typical scale of 1 to 10. This point was illustrated by asking subjects to assign numbers to a series of distances the experimenter delimited between his or her hands. Subjects who initially had difficulty understanding the concept of ratio estimation were briefly coached during this practice period (e.g. 'Is this distance about twice as great as the previous distance?'). During the experiment subjects entered their numerical estimates on a computer keyboard.

Taste scaling

Subjects. Eighteen subjects (nine men and nine women, median age = 26 years) participated in the taste scaling portion of the experiment. None had prior experience with either scaling method.

Stimuli. The taste stimuli were NaCl and sucrose prepared in the same five concentrations (0.15, 0.26, 0.47, 0.84 and 1.5 M) in deionized H_2O. Sucrose and NaCl were chosen as before because they are considered prototypic gustatory stimuli, yet NaCl was known to cause sensory irritation as well as taste at moderate and high concentrations (Abrahams *et al.*, 1937; Holway and Hurvich, 1937; Green and Gelhard, 1989). The stimuli were sampled in 10-ml aliquots individually poured into 1-oz medicine cups.

Psychophysical procedure. The procedure was the same as that used for scaling odors with the exception of some details related to stimulus delivery and perceptual context. Specifically, the stimuli were presented three times each rather than twice and the ISI was approximately 1 min rather than 30 s. On each trial the subject sipped the sample, expectorated it after 3 s, rated sensation intensity and then rinsed at least twice before the next stimulus was presented.

The following instructions were read to establish the perceptual context for the ratings of taste intensity:

> In making your judgements of taste, you should rate the stimuli relative to other tastes of all kinds that you have experienced. Thus, 'strongest imaginable' refers to the most intense sensation of taste that you can ever imagine experiencing. This includes such varied sensations as those produced by a fresh lemon, a piece of celery, or spicy mustard. Note that by 'taste' we do not mean the pain produced by a physical trauma like biting or burning your tongue. Simply rate the samples relative to tastes that you experience in daily life.

The instructions and training for magnitude estimation were the same as in the odor scaling task.

Data analysis. Prior to analysis, the data for both methods were averaged over replicates to yield a best estimate of perceived intensity for each stimulus. Because ME data tend to be skewed in a manner approximating a log-normal distribution (Marks, 1974) and because visual evaluation of data from the LMS have shown a similar distribution across subjects (Green *et al.*, 1993), the raw data were normalized by converting to logarithms. Subsequent operations on the data sets were performed on the logarithms of the means.

Before comparing the magnitude estimates to the scale values, the irrelevant differences among subjects created by the free use of numbers in magnitude estimation were eliminated by equating each individual's grand mean of the

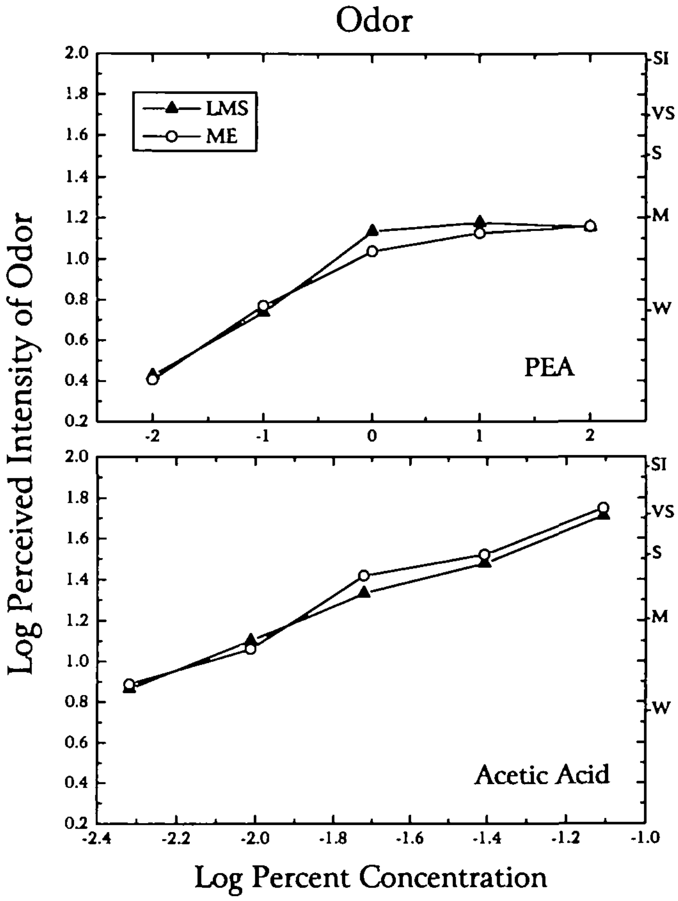

Figure 2 The nearly identical psychophysical functions derived from the ME and LMS tasks are shown for the two odor stimuli of Experiment 1. To enable a direct comparison of the methods, the ME data have been standardized to the $\log_{10}$ of the LMS data. The letters on the right y-axes represent verbal labels of the LMS: W = weak, M = moderate, S = strong, VS = very strong, SI = strongest imaginable (barely detectable was omitted because it falls below 0.2). Thus, on average, the three highest concentrations of PEA produced only moderate odors, whereas the highest concentration of acetic acid produced a very strong odor.

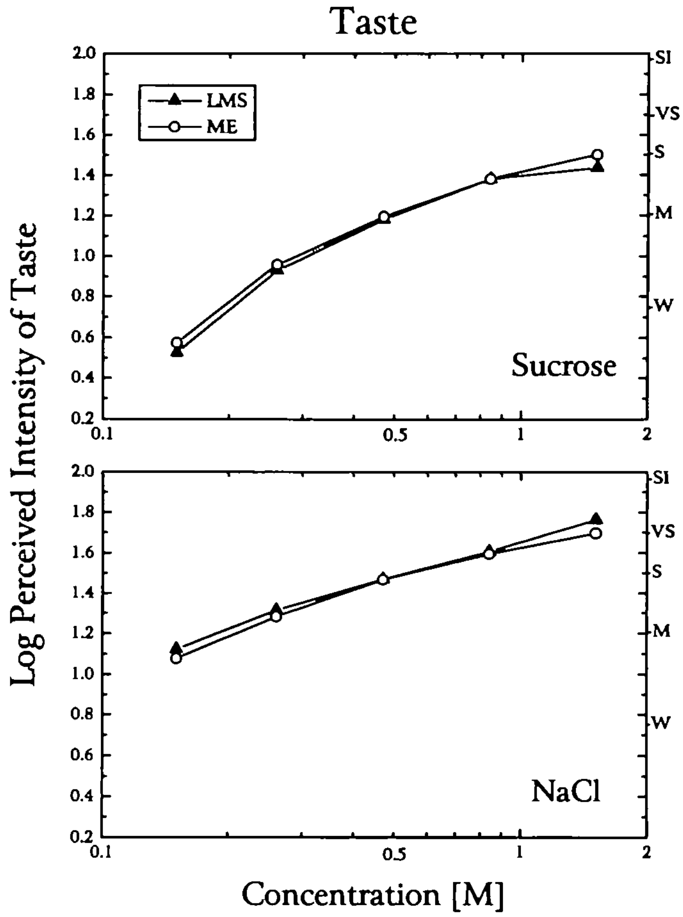

Figure 3 The same as Figure 2, except for the two taste stimuli of Experiment 1. Note that the highest sucrose concentration produced, on average, a little less than a strong taste, whereas the highest concentration of NaCl produced a very strong taste

logarithms for ME to the grand mean of the logarithms for the LMS values. This avoided the creation of a difference between methods based solely upon number usage. Because this manipulation also reduced the between-subject variance in the magnitude estimation condition, individual subject means in the scale condition were also normalized to the grand (log) mean of the scale. Repeated measures ANOVAs were conducted on the log values to assess possible effects of and interactions among three factors: scaling method, concentration and chemical.

Results

Odor

The data from the odor intensity task are shown in Figure 2. The normalized mean ratings for PEA and acetic acid were similar for the two scaling methods. Although there were main effects of stimulus [$F(1,17) = 25.8$, $P < 0.0001$] and concentration [$F(4,68) = 126.6$, $P < 0.0001$] and an interaction between these two variables [$F(4,68) = 3.4$, $P < 0.025$], there was not a significant interaction between method and concentration [$F(4,68) = 0.02$, $P > 0.05$], or among stimulus, concentration and method [$F(4,68) = 0.63$, $P > 0.05$]. The absence of the three-way interaction confirms the visual impression (Figure 2) that the two methods produced nearly identical psychophysical functions for both chemicals. The similarity between methods is particularly striking given the different shapes of the psychophysical functions for PEA and acetic acid. The PEA data are well described by a two-limb power function that has slopes of approximately 0.3 (LMS = 0.35, ME = 0.31) and 0 (LMS = 0.01, ME = 0.06). The slope of zero over the higher concentrations implies subjects were unable to perceive a difference among those concentrations. In contrast, the acetic acid data are well described by a single power function having a slope of approximately 0.7 (LMS = 0.69, ME = 0.72).

Taste

Consistent with the outcome for odor stimuli, Figure 3 indicates that the psychophysical functions for taste produced by the two scaling methods were nearly identical. The ANOVA indicated significant main effects of stimulus [$F(1,17) = 79.0$, $P < 0.0001$] and concentration [$F(4,68) = 95.4$, $P < 0.0001$] and a significant interaction between stimulus and concentration [$F(4,68) = 11.3$, $P < 0.0001$]. The latter interaction derives from the different shapes of the functions for sweetness and saltiness. There was neither a significant two-way interaction between method and concentration [$F(4,68) = 0.04$, $P > 0.05$] nor a three-way interaction between method, stimulus and concentration [$F(4,68) = 0.49$, $P > 0.05$], which indicates the two methods produced comparable psychophysical functions. The statistical findings are borne out by the slopes of the best-fitting power functions: for sucrose the slopes are 0.63 for the LMS and 0.62 for ME; for NaCl the methods produced identical slopes of 0.90.

Experiment 2: individual taste qualities

The results of Experiment 1 indicated the LMS could be used as a scale of smell intensity or taste intensity, and still yield data equivalent to that produced by magnitude estimation. However, because sensations had been rated within the context of all possible odors or tastes, the results did not allow inferences to be made about sensation intensity within a particular odor or taste quality (e.g. sweetness). 'Absolute' information about sensation intensity within a specific taste quality can be acquired only if judgements of intensity are made within the context of that quality alone (e.g. relative to the strongest imaginable sweetness). The same would hold for evaluating specific odor qualities, although, unlike taste, the seemingly infinite variety of olfactory sensations precludes classification of odors into a small set of universally agreed upon qualities.

The present experiment was therefore designed to discover whether defining the top of the LMS as the strongest imaginable taste of a particular quality would produce data comparable to that produced by ME. The qualities of sweetness, saltiness and bitterness were studied.

Method

Subjects

Three groups of adults (two groups of 20 and a third group of 17) were paid to take part in the experiment. None of the subjects had prior experience with either the LMS or ME. One group was tested with sucrose (10 males and 10 females, median age = 21 years), another with NaCl (10 males and 10 females, median age = 23 years) and a third with quinine sulfate (eight females and nine males, median age = 26 years). Twenty subjects initially began testing in the quinine experiment, but three were dropped from the analysis after it was discovered they had consistently reported 'no sensation' from two or more of the concentrations in the series.

Stimuli

The stimuli were 10-ml samples of six concentrations of sucrose (0.056, 0.10, 0.18, 0.32, 0.56, 1.0 M), NaCl (0.032, 0.056, 0.10, 0.18, 0.32, 0.56 M) or QSO_4 (0.32, 0.18, 0.10, 0.056, 0.032, 0.018 mM) dissolved in dH_2O. Sucrose and NaCl were chosen as before because they are both prototypical gustatory stimuli and NaCl stimulates the trigeminal system at high concentrations; QSO_4 was chosen because intense bitterness may represent the 'strongest imaginable' purely gustatory sensation.

Psychophysical procedure

The three taste qualities were studied sequentially over a period of several months in the order sweetness, saltiness and bitterness. The general experimental design was the same as had been used for the taste stimuli in Experiment 1. Subjects served in two sessions, one in which they used the LMS and one in which they used ME. Half of the subjects received the LMS first and the other half received ME first. Each concentration was presented three times per session in random sequence. On each trial the subject was prompted to sip the sample, expectorate after 3 s, rate sensation intensity, then to rinse at least twice before the next stimulus was presented 1 min later.

The instructions for using the LMS were changed from Experiment 1 by eliminating reference to taste, in general, and by asking subjects to rate sensation intensity in terms of the strongest imaginable sweetness, saltiness or bitterness. The portion of the LMS instructions relevant to the context in which judgements were to be made was therefore reduced to the following:

In making your judgements of intensity you should rate the stimuli relative to the strongest sensation of sweetness [saltiness, bitterness] you have experienced. Thus, 'strongest imaginable' refers to the most intense sensation of sweetness [saltiness, bitterness] that you can ever imagine experiencing.

The instructions for ME were the same as had been used for the taste portion of Experiment 1, except that subjects were told the sensations would be sweet, salty or bitter, depending upon the experiment.

Data analysis

The data were treated in the same manner as Experiment 1, except that the results for the three taste qualities were analysed in separate ANOVAs.

Results

The results for the three taste qualities are shown in Figure 4. It is clear from the figure that in no case did the two scaling methods yield identical psychophysical functions. The slopes of the best-fitting power functions produced by the LMS were uniformly steeper than those produced by ME (33% steeper for sweetness and saltiness, 23% steeper for bitterness). Statistical interactions between method and concentration [for sweetness, $F(5,95) = 7.0$, $P < 0.0001$; for saltiness, $F(5,95) = 6.9$, $P < 0.0001$; for bitterness, $F(5,85) = 3.1$, $P < 0.0136$)] confirmed that the observed differences were significant. It is unclear why the functions for bitterness deviated so much from linearity on log-log co-ordinates, although individual differences in sensitivity to QSO_4 seemed to be a contributing factor. In addition to the three subjects who were dropped from the analysis because they repeatedly reported no bitterness at the lower concentrations, some of the remaining subjects also failed to report bitterness at those concentrations on at least one trial.

Overall, these results indicate the LMS cannot be assumed to yield data equivalent to ME or to have ratio-level properties, if ratings of taste intensity are made within the context of a specific taste quality. Although a sour stimulus was not tested, the uniform disagreement with ME across three stimuli having such varied sensorineural and perceptual effects constitutes strong evidence for the generalizability of the finding to other taste stimuli.

It should be pointed out that forcing the grand mean of the ME data to equal the grand mean of the LMS data dictated that a difference in slope between methods would result in the crossing interaction we observed. Consequently, it cannot be concluded, as the data in Figure 4 seem to suggest, that compared to ME the LMS produced lower intensity ratings at low concentrations and higher intensity ratings at high concentrations. We can only conclude that compared to ME, the LMS indicated a more rapid rate of change in perceived intensity across concentrations.

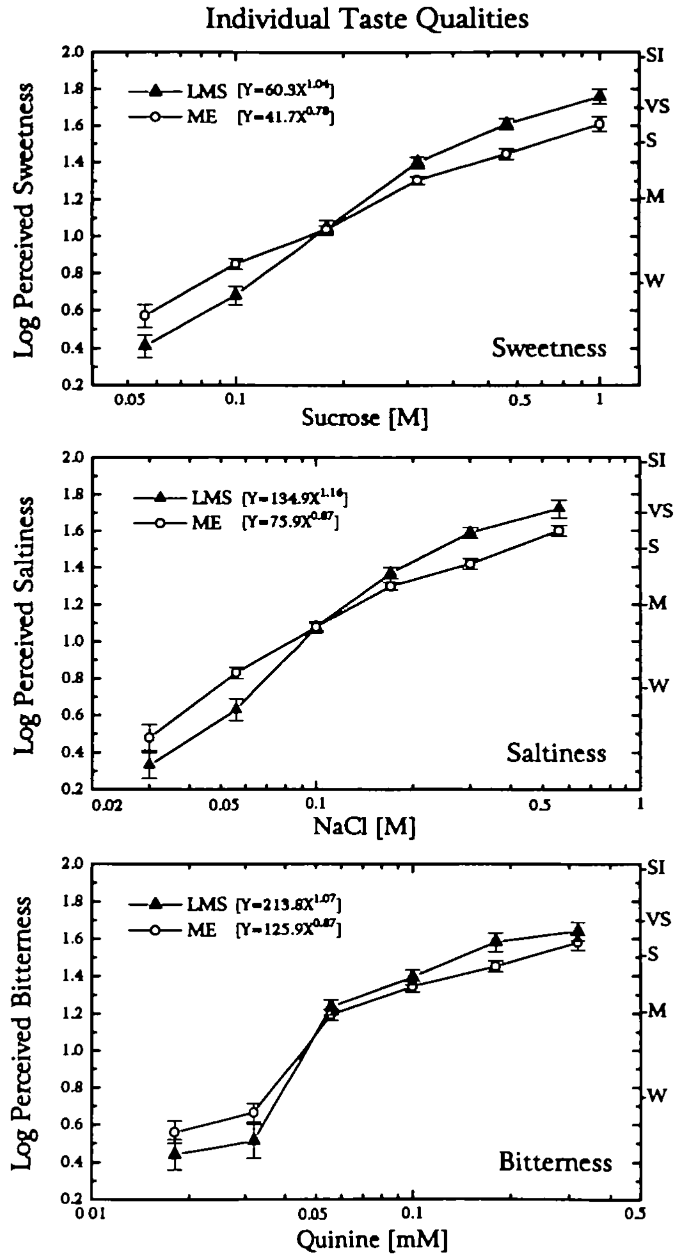

Figure 4 Same as Figures 2 and 3, except for the individual taste qualities of Experiment 2. Vertical bars represent standard error (SE) of the standardized $\log_{10}$ means. Note that the semantic information is valid only for the LMS data and, because the ME data have in each case been standardized to the grand $\log_{10}$ mean of the LMS data, the relative positions of the ME and LMS functions have been artificially fixed.

Experiment 3: taste excluding pain

The disagreement between the results from the two scaling methods in Experiment 2 implies that the range and/or form of the LMS is inappropriate for scaling the intensity of specific taste qualities. Yet, the two methods produced nearly identical psychophysical functions in Experiment 1, when intensity was scaled within the context of all taste sensations. The most parsimonious explanation for these two outcomes is that the perceptual range of 'all taste sensations' is more similar to the perceptual range of 'all oral sensations' than is the perceptual range of a single taste quality. The instructions to subjects in Experiment 1 were consistent with this interpretation: by including the 'taste' of spicy mustard as an exemplar, subjects were encouraged to adopt a definition of taste that included chemesthetic sensations such as burning and stinging. Although irritants in foods are unlikely to produce sensations rivaling the strongest oral pains, they can, nevertheless, provoke frankly painful sensations (e.g. the burn of a hot pepper). Accordingly, it is likely that in Experiment 1 the 'strongest imaginable' sensations of both taste and smell (for which irritating exemplars were also given) were assumed to be painful, whereas in Experiment 2 pain was eliminated from consideration when subjects were told that the top of the scale was the strongest imaginable sweetness, saltiness or bitterness.

The present experiment, therefore, provided a test of the hypothesis that the agreement between methods in Experiment 1 was attributable to the inclusion of potentially painful chemesthetic sensations within the concept of taste (and smell). The subjects' task in the experiment was to rate taste sensations produced by sucrose within the context of all taste sensations except painful or irritating ones (i.e. sweet, sour, salty and bitter).

Method

Subjects

Eighteen people (13 women and five men, average age = 24.8 years) were paid to serve in the experiment. As before, none were experienced with the LMS or ME.

Stimuli

The stimuli were seven concentrations of sucrose in 1/4 log steps (0.063, 0.100, 0.16, 0.25, 0.40, 0.63 and 1.0 M). Sucrose was chosen, as in Experiment 1, because it is the prototypical non-irritating gustatory stimulus and, hence, the best candidate for evaluating taste intensity independently

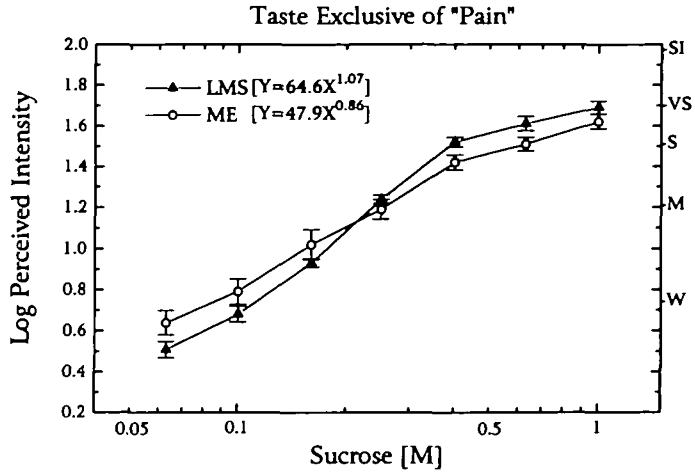

Figure 5 The results from the LMS and ME when the instructions for use of the LMS were to rate sensations produced by sucrose in the context of 'taste exclusive of pain' (Experiment 3). The data were standardized as before and the vertical bars again represent SE.

of irritation or pain. The method of stimulus preparation and delivery was the same as before.

Psychophysical procedure

The temporal parameters of stimulus delivery and intensity scaling were also the same as before. Subjects again served in two sessions, one in which they used the LMS and one in which they used ME, with half receiving the LMS first and half receiving ME first. The key manipulation in the experiment was the instruction to subjects to exclude sensations of irritation and pain from the concept of 'taste' when using the LMS. The relevant portion of the LMS instructions, with the changes in wording from Experiment 1 highlighted in italics, was as follows:

> Thus, 'strongest imaginable' refers to the most intense sensation of *sweetness, bitterness, sourness or saltiness* that you can ever imagine experiencing. Note that by 'taste' we do not mean the *heat or pain produced by hot and spicy foods.* Simply rate the samples relative to tastes that you experience in daily life.

Along with these differences, no examples were given of tastes that have chemesthetic qualities. The instructions for ME were the same as in the previous experiments.

Data analysis

The data were logged and standardized as before, and analysed using a repeated measures ANOVA.

Results

The results shown in Figure 5 are consistent with the hypothesis that the LMS does not produce data comparable to ME if the sensory dimension being scaled does not include painful sensations. As in Experiment 2, there was a significant method × concentration interaction [$F(6,102) = 5.51$, $P < 0.0001$], which reflected a difference in the slopes of the psychophysical functions produced by the two methods. The exponents of the best-fitting power functions were 1.07 for the LMS and 0.86 for ME, a difference of 24.4%, which falls close to the 23% difference obtained for bitterness but below the 33% difference in slopes obtained for sweetness and saltiness in Experiment 2.

Discussion

The LMS as a scale of taste and smell

The present study has demonstrated that with appropriate instructions the LMS can be used to measure the intensity of sensations of both taste and smell. This outcome extends the findings of the previous study of the LMS in which the psychophysical functions for gustatory, chemesthetic and thermal stimuli were found to agree closely with the functions produced by ME when the three forms of stimulation were rated relative to the strongest imaginable oral sensation (Green *et al.*, 1993). The agreement with ME in the earlier study implied the LMS could be used with confidence to determine the slopes of psychophysical functions for gustatory as well as for somesthetic stimuli. However, the perceptual context in which the intensity ratings had been made permitted interpretation of the semantic information from the LMS only in terms of 'oral sensation'. This is an important point, because researchers in the chemical senses often wish to measure the intensity of tastes or smells *per se*. To this end, Experiment 1 demonstrated that equivalence with ME could be maintained if taste intensity or smell intensity were rated relative to the strongest imaginable taste or smell. This means that in theory the LMS can be used both to determine the relative intensities of different tastes or smells on a ratio scale and to provide semantic information about their 'absolute' intensities within each perceptual domain. Moreover, the use in Experiment 1 of chemicals (NaCl and acetic acid) that cause irritation at high concentrations provides evidence that the LMS is as effective as ME for revealing differences in the slope and form of psychophysical functions caused by differences in sensory processing.

Limitations with specific taste qualities

The results of Experiments 2 and 3 place constraints on the ways in which the scale can be used to measure gustatory sensations, if equivalence with ME and the assumption of ratio level data are to be maintained. In Experiment 2 the steeper psychophysical functions produced by the LMS in the context of a specific taste quality led to the hypothesis that this disparity was attributable to the absence of painful sensations among the strongest imaginable sensations of sweetness, saltiness and bitterness. The data from Experiment 3, in which the instructions to subjects explicitly eliminated painful sensations from the concept of 'taste', support this hypothesis. Also compatible with the hypothesis is the semantic information about sucrose stimulation obtained in the different experiments. Whereas a 1.5 M solution of sucrose produced a mean response just below 'strong' on a scale of all 'tastes' in Experiment 1, a 1.0 M solution produced mean responses at and above 'very strong' on scales of 'sweetness' and 'taste excluding pain' in Experiments 2 and 3, respectively. These shifts in the judged 'absolute' magnitude of sucrose stimulation are consistent with a more intense 'strongest imaginable' sensation in Experiment 1 than in the other two experiments, and their occurrence provides clear evidence that the instructions pertaining to perceptual range affect how the scale is used and how its data should be interpreted. Overall, the data from Experiments 2 and 3 indicate that the LMS needs to be modified if it is to be used to obtain data equivalent to ME on sensory continua—like those of sweet, sour, salty and bitter—that do not include pain. Whether a single scale will accommodate all such continua will have to be determined empirically. However, a need for multiple versions of the LMS may be hinted by the fact that the difference in slopes of the psychophysical functions between methods varied from 33% for sweetness and saltiness, to 23–24% for bitterness and taste exclusive of pain.

The same limitation is likely to be encountered in scaling odors that do not produce painful or irritating sensations at high concentrations. Thus, for example, if the intent is to scale sensations of 'rose odor' or 'vanilla odor', the LMS is probably inappropriate. As noted earlier, however, the much greater variety of olfactory qualities makes it difficult to arrive at meaningful qualitative categories analogous to sweetness or saltiness, and as a result odor intensity is usually rated independently of quality. The results of Experiment 1 indicate that use of the LMS in this more general way produces data comparable to ME.

Implications for use in other sensory modalities

Although the results raise caveats about using the LMS to scale intensity within specific taste or odor qualities, they also imply that it may be appropriate for use in the variety of perceptual domains that include pain and perhaps for pain itself. For example, inasmuch as sounds can be painfully loud and lights painfully bright, the LMS might be used successfully to scale loudness or brightness. Marks *et al.* (1983) used Borg's CR scale to measure perceived loudness and as judged by comparisons with psychophysical exponents obtained with ME, had only limited success. Because both the range and distribution of the semantic labels differ somewhat between the LMS and CR scales (Green *et al.*, 1993), greater success might be expected with the LMS. However, as in the present study, the success of the LMS in these domains may well depend on the instructions subjects receive. Unless it is made clear that painfully intense sounds and lights are to be included in the concept of 'strongest imaginable', some or all subjects may limit the judged perceptual range to non-painful sensations and produce data similar to those obtained for 'taste exclusive of pain' (Experiment 3). The same problem would not, of course, be encountered in the measurement of pain itself. With an upper bound essentially equivalent to the strongest imaginable oral pain, the LMS might be particularly well suited for scaling pain. The semantic scaling data of Tursky (1976) and Gracely *et al.* (1978a,b) support this view. Tursky reported the range between semantic magnitudes of 'just noticeable' and 'excruciating' pain to be 1.45 log units, and Gracely *et al.* reported semantic magnitudes for 'extremely weak' and 'extremely intense' pain that differed by between 1.8 and 1.9 log units. The LMS compares favorably with a range between 'barely detectable' and 'strongest imaginable' of 1.84 log units.

While the LMS may eventually prove useful in a number of perceptual domains, the lack of equivalence with ME for specific taste qualities conflicts with the idea that the perceptual range is the same in all sensory modalities (Teghtsoonian, 1971, 1973; Borg, 1982). The steeper slopes obtained with the LMS compared to ME for ratings of taste *per se* and specific taste qualities suggests that, consistent with the absence of pain on the taste continuum, the perceptual range of taste is smaller than the perceptual range of somesthesis. We would therefore predict that if the scale-construction task (Green *et al.*, 1993) were repeated within the context of a specific taste quality, the ratio between the mean magnitude estimates of 'barely detectable' and 'strongest imaginable' sensations would be significantly less, thereby making the response range smaller than on the current LMS. It would be interesting to learn whether the spacing among the other semantic labels might also differ for a scale specific to a single taste quality. It is possible, for example, that 'moderate' would lie closer to the middle of the perceptual range if the top of the range were no longer defined by pain.

The relationship between the LMS and ME

It deserves emphasis that we compared the LMS to ME because ME is believed to yield ratio-level data on perceived intensity (Stevens and Galanter, 1957; Stevens, 1971; Marks, 1974). Although a critical evaluation of this contention is beyond the scope of the present paper, it may be instructive to consider the relationship between numerical ratio estimation and the semantically-based task of the LMS. As was pointed out in the introduction, Borg and his colleagues (Borg, 1982; Borg *et al.*, 1985; Borg and Borg, 1987; Marks *et al.*, 1983, 1992) worked from the assumption that the key to obtaining ratio-level data lies in determining the appropriate spacing among a set of verbal labels that spans the full range of conceivable perceptual intensities and that this spacing can be accomplished via a semantic scaling task. If the obtained semantic magnitudes are valid, the resulting scale should provide a linear transformation of the internal representation of sensation intensity in the modality of interest. By comparison, a typical category scale, on which semantic labels are arbitrarily spaced at even intervals along the scale, is thought to provide a non-linear transformation of perceived intensity on prothetic (Stevens and Galanter, 1957) continua, as evidenced by the curvilinear relationship typically found when data from ME are plotted against data from category scales (Stevens and Galanter, 1957; Marks, 1968; Stevens, 1971; Gibson and Tomko, 1972). The location of 'moderate' in the bottom one-fifth of the LMS rather than in the middle graphically illustrates a similar non-linear relationship between equal-interval category scales and the LMS (e.g. Ellermeier *et al.*, 1991).

Another way to understand the relationship between the LMS and ME is to consider that the scale substitutes the task of matching numbers to sensations with the task of matching verbal labels to sensations. Although the use of verbal labels to obtain quantitative data on sensation magnitude might at first appear less direct and, hence, perhaps less valid than numerical ratio estimation, the opposite argument can be made: because in daily life the strength of sensations is more often communicated with words (e.g. a 'strong' smell, a 'weak' cup of coffee) than with numbers, a task that requires responses based on words may be both more natural and more direct than a task based on numbers. Indeed, although in theory the aim of ME is to match numbers to sensations (Stevens, 1971), subjects are typically instructed to estimate the ratios among sensations (e.g. 'if a sensation is twice as strong, assign a number twice as large'), rather than to assign numbers directly to sensations. Omitting such instructions risks the possibility subjects will limit their responses to the familiar, categorical numerical range of 1–10 (see, however, Zwislocki, 1980; Gescheider and Hughson, 1991). Ratio estimation can therefore be viewed as a strategy for teaching subjects how to do something they rarely, if ever, attempt: to make one-to-one matches between numbers and sensations. Presumably, the extent to which this strategy succeeds determines the quality of the numerical matches and the validity of the resulting psychophysical relationship. In contrast, the present results and those from the previous study of the LMS (Green *et al.*, 1993) suggest that defining the proper perceptual context enables subjects to find locations on a familiar semantic continuum that 'match' the perceptual intensities. Thus, the LMS and other empirically determined CR scales can be thought of as natural alternatives to ME that provide the same information about perceptual magnitude while also offering semantic information about the (absolute) intensity of sensations within the perceptual context of interest. It will eventually be important to determine how changes in stimulus variables that typically induce strong contextual effects (Parducci and Perrett, 1971; Poulton, 1979; Mellers and Birnbaum, 1982; Algom and Marks, 1990; Marks, 1991; Rankin and Marks, 1991) influence responses obtained using the LMS. The central question is whether the perceptual frame subjects bring to the laboratory, and which they are instructed to use with the LMS, persists in the face of varying stimulus contexts.

Summary

The results of this study indicate the LMS is a valid alternative to ME as a tool for measuring the perceived intensity of gustatory, olfactory and chemesthetic sensations within the broadly defined perceptual domains of taste and smell. The present findings together with those from a previous study of the LMS (Green *et al.*, 1993) indicate that divergence from ME occurs only when subjects are instructed to rate sensations on the LMS within the context of specific taste qualities or taste sensations exclusive of pain. This pattern of results implies that the response range of the LMS is too large for continua that do not include pain. Hence, additional semantic scaling is called for to discover how the range and distribution of semantic descriptors on the LMS should be modified to accommodate the smaller perceptual range (or ranges) of such continua.

ACKNOWLEDGEMENTS

The authors thank John Pierce for his contributions to some of the early discussions and data collection in Experiment 1. This research was supported in part by NIH grants DC00249–11 and DC02499–01.

REFERENCES

Abrahams, H., Krakauer, D. and Dallenbach, K.M. (1937) Gustatory adaptation to salt. *Am. J. Psychol.*, **49**, 462–469.

Algom, D. and Marks, L.E. (1990) Range and regression, loudness scales, and loudness processing: toward a context-bound psychophysics. *J. Exp. Psychol.: Hum. Percept. Perform.*, **16**, 706–727.

Borg, G. (1982) A category scale with ratio properties for intermodal and interindividual comparisons. In Geissler, H-G. and Petxold, P. (eds), *Psychophysical Judgement and the Process of Perception*. VEB Deutxcher Verlag der Wissenschaften, Berlin, pp. 25–34.

Borg, G. (1990) Psychophysical scaling with applications in physical work and the perception of exertion. *Scand. J. Work Environ. Hlth*, **16**, 55–58.

Borg, G. and Borg, P. (1987) On the relations between category scales and ratio scales and a method for scale transformation. *Reprints from the Department of Psychol., University of Stockholm*, **672**, 1–14.

Borg, G., Ljunggren, G. and Marks, L.E. (1985) General differential aspects of perceived exertion and loudness assessed by two new methods. *Reprints from the Department of Psychol., University of Stockholm*, **636**, 1–13.

Doty, R.L., Brugger, W.E., Jurs, P.C., Orndorff, M.A., Snyder, P.J. and Lowry, L.D. (1978) Intranasal trigeminal stimulation from odorous volatiles: Psychometric responses from anosmic and normal humans. *Physiol. Behav.*, **20,** 175–185.

Ellermeier, W., Westphal, W. and Heidenfelder, M. (1991) On the 'absoluteness' of category and magnitude scales of pain. *Percept. Psychophys.*, **49,** 159–166.

Gescheider, G.A. and Hughson, B.A. (1991) Stimulus context and absolute magnitude estimation: a study of individual differences. *Percept. Psychophys.*, **50,** 45–57.

Gibson, R.H. and Tomko, D.L. (1972) The relation between category and magnitude estimates of tactile intensity. *Percept. Psychophys.*, **12,** 135–138.

Gracely, R.H., McGrath, P. and Dubner, R. (1978a) Ratio scales of sensory and affective verbal pain descriptors. *Pain*, **5,** 5–18.

Gracely, R.H., McGrath, P. and Dubner, R. (1978b) Validity and sensitivity of ratio scales of sensory and affective verbal pain descriptors: manipulation of affect by diazepam. *Pain*, **5,** 19–29.

Green, B.G. and Gelhard, B. (1989) Salt as an oral irritant. *Chem. Senses*, **14,** 259–271.

Green, B.G., Shaffer, G.S. and Gilmore, M.M. (1993) Derivation and evaluation of a semantic scale of oral sensation magnitude with apparent ratio properties. *Chem. Senses*, **18,** 683–702.

Holway, A.H. and Hurvich, L.M. (1937) Differential gustatory sensitivity to salt. *Am. J. Psychol.*, **48,** 49–37.

Kobal, G. and Hummel, C. (1988) Cerebral chemosensory evoked potentials elicited by chemical stimulation of the human olfactory and respiratory nasal mucosa. *Electroencephalogr. Clin. Neurophysiol.*, **71,** 241–250.

Marks, L.E. (1968) Stimulus-range, number of categories, and form of the category-scale. *Am. J. Psychol.*, **81,** 467–479.

Marks, L.E. (1974) *Sensory Processes*. Academic Press, New York.

Marks, L.E. (1991) Reliability of magnitude matching. *Percept. Psychophys.*, **49,** 31–37.

Marks, L.E., Borg, G. and Ljunggren, G. (1983) Individual differences in perceived exertion assessed by two new methods. *Percept. Psychophys.*, **34,** 280–288.

Marks, L.E., Borg, G. and Westerlund, J. (1992) Differences in taste perception assessed by magnitude matching and by category-ratio scaling. *Chem. Senses*, **17,** 493–506.

Mellers, B.A. and Birnbaum, M.H. (1982) Loci of contextual effects in judgment. *J. Exp. Psychol.: Hum. Percept. Perform.*, **111,** 242–270.

Parducci, A. and Perrett, L.F. (1971) Category rating scales: Effects of relative spacing and frequency of stimulus values. *J. Exp. Psychol. Monogr.*, **89,** 427–452.

Poulton, E.C. (1979) Models for biases in judging sensory magnitude. *Psycholog. Bull.*, **86,** 777–803.

Rankin, K.M. and Marks, L.E. (1991) Differential context effects in taste perception. *Chem. Senses*, **16,** 617–629.

Stevens, S.S. (1971) Issues in psychophysical measurement. *Psycholog. Rev.*, **78,** 426–450.

Stevens, S.S. and Galanter, E.H. (1957) Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.*, **54,** 377–411.

Teghtsoonian, R. (1971) On the exponents in Stevens' law and the constant in Ekman's law. *Psycholog. Rev.*, **78,** 71–80.

Teghtsoonian, R. (1973) Range effects in psychophysical scaling and a revision of Stevens' law. *Am. J. Psychol.*, **86,** 3–27.

Tursky, B. (1976) The development of a pain perception profile· A psychophysical approach. In Weisenberg, M. and Tursky, B.(eds), *New Perspectives in Therapy and Research.* Plenum Press, New York, pp. 171–194.

Zwislocki, J.J. (1980) Absolute and other scales: question of validity. *Percept. Psychophys.*, **33,** 593–594.

Received on August 28, 1995; accepted on January 29, 1996