Measurement of Pain Using the Visual Numeric Scale

PHILIP L. RITTER, VIRGINIA M. GONZÁLEZ, DIANA D. LAURENT, and KATE R. LORIG

ABSTRACT. ***Objective***. We introduce the English-language Visual Numeric Scale (VNS) for self-reported pain and examine its psychometric properties; we compare the VNS to the better known Visual Analog Scale (VAS).

Methods. We developed the VNS, which combines strong visual cues with an 11-point numeric rating scale. The VNS was administered to 2 sets of subjects with arthritis or chronic disease (N = 175, N = 192, respectively) and responses were examined. To compare the VNS to the VAS, we administered both scales to all subjects and used correlations to compare them to each other and to health distress and overall general health scores. A subset of respondents enrolled in an arthritis self-management program were given the VNS 4 months later, and change scores were used to test the sensitivity of the VNS.

Results. The VNS had means of 5.4 and 5.6 in the 2 samples, with distributions across the range of possible values. The VNS correlated well with the VAS (r = 0.85) and correlated slightly better than the VAS with the 2 independent health measures. The VNS was more likely to be completed than the VAS and there were fewer coding errors with the VNS. The VNS showed a significant (effect size 0.28) positive change for participants in a self-management course.

Conclusion. The VNS appeared to be a valid measure. It was as successful as the VAS in measuring the underlying pain variable. It was easier to administer and code than the VAS, and was sensitive to change in pain. (J Rheumatol 2006;33:574–80)

Key Indexing Terms:
PAIN MEASUREMENT PSYCHOMETRICS EVALUATION

Pain is a key outcome for many arthritis studies. Its measurement is of importance for those conducting clinical trials or outcomes research. We examined the properties of the recently developed Visual Numeric Scale (VNS) and compared it to the more traditional Visual Analog Scale (VAS). The VNS combines the features of a numeric scale with strong visual clues, including size and shading.

In a summary of self-report pain scales, Jensen and Karoly[1] describe the 3 most commonly used single-item methods to measure pain intensity. (1) Verbal rating scales (VRS) consist of a series of descriptive words (e.g., "no pain" to "severe pain"). (2) Visual analog scales generally consist of a 10 centimeter line with anchor terms at each end, e.g., "No pain" or "Lack of any pain" at one end versus "Severe pain" or "Pain as bad as it could be" at the other end. Subjects may mark any point along the line and scoring may range from the nearest centimeter (0–10) to the nearest millimeter (0–100). (3) Numeric rating scales ask subjects to rate their pain using numbers, usually 0 to 10 (11 points) or 0 to 20 (21 points). Self-administered numeric scales may include anchor terms similar to the VAS.

From the Stanford University School of Medicine, Palo Alto, California, USA.

Supported in part by grant No. 1-R01-NR-03146-01 from the National Institute for Nursing Research and by the American College of Rheumatology.

P.L. Ritter, PhD; V.M. González, MPH; D.D. Laurent, MPH; K.R. Lorig, DPH.

Address reprint requests to P.L. Ritter, Stanford Patient Education Research Center, 1000 Welch Road, Suite 204, Palo Alto, CA 94034. E-mail: philr@stanford.edu

Accepted for publication October 28, 2005.

VRS are relatively easy to administer and score for literate subjects, but may cause difficulties for the less literate or those where the language of the scale is not the first language. Unless researchers resort to the relatively complicated cross-modality matching procedures[2], one cannot assume equal intervals between responses. VAS may have problems with subjects understanding the scale without careful supervision[3], and scoring may also introduce errors[4]. A major advantage is that VAS tend to approximate ratio-level scales for groups of people[5]. Numeric scales are more easily scored than VAS and can also be administered verbally (i.e., in followup data collection by telephone). Direct comparisons of numeric scales with VAS have tended to support use of numeric scales, especially for self-administered questionnaires[4,6,7].

A number of other single-item scales have been developed and tested. A series of faces (from sad or crying to smiling) has been used successfully with young children[8]. There have been several scales that combine visual, numeric, and/or verbal cues. The original inspiration for the VAS, referred to as the Graphic Rating Scales, included descriptive words along the line (such as "mild," "moderate," and "severe")[9]. Scott and Huskisson found that adding descriptive terms along the VAS resulted in strong clustering near the points closest to the terms. Thus a more pure form (i.e., without descriptive terms) of the VAS was preferable[3]. In contrast, a later study supported the use of such a scale, the Word Graphic Scale, among children and adolescents[10]. A method used by McCaffery and Beebe[11] incorporates numbers, descriptive words, and colors along a VAS. A different colored analog scale designed to measure children's pain was compared to the VAS, and found

to have appropriate psychophysical properties and discriminant validity[12]. Similar "pain thermometers" or "pain rulers" are often used in clinical contexts.

In a direct comparison of 0–10 single-item pain intensity ratings with multi-item composite scales, Jensen, *et al*[13] found that the single-item measures were as sensitive as multi-item measures. They concluded that although multi-item composite scales may be more appropriate in a clinical setting and with small groups because of higher reliability, individual 0–10 pain intensity ratings have sufficient psychometric strengths to be used in chronic pain research, especially research involving group comparisons with relatively large sample size.

The Visual Numeric Scale was developed to take advantage of the features of numeric scales while providing multiple visual cues. These include height and shading of bars associated with each numeral (Figure 1), and can thus be considered a combination scale with both visual and numeric components. It was developed after we experienced strong negative feedback regarding the VAS from leaders and focus group participants involved in a Spanish-language arthritis patient-education program. Our goal was a single-item pain scale that could be used unassisted by adults in self-report questionnaires and that would be appropriate for comparisons of change scores for groups of participants in self-management programs. We also needed to avoid the expense of color printing. As Champion, *et al*[14] and others have pointed out, no scale is globally valid or invalid; each must be judged in relation to a specific purpose. The Spanish version of the VNS met our criteria and proved easy to understand and administer, and was successful in measuring differences in pain[15,16]. A comparison of the Spanish VNS with a Spanish VAS found a correlation of 0.72, with 24% missing data for the VAS but only 6% for the VNS[15,16]. A description of the Spanish VAS was also included in a book on outcome measures for health education[17]. Subsequently, we developed an English version of the VNS for use in evaluating patient education programs among English speaking patients. A copy of the English VNS may be downloaded from the Stanford Patient Education Website at http://patienteducation.stanford.edu/research/vnspain.pdf.

Although the Spanish VNS proved relatively easy to understand for participants filling out questionnaires and had good reliability and construct validity[15], one can legitimately ask how well the English-language VNS meets the same criteria: how well does it consistently measure underlying subjective pain and how sensitive is it to changes in pain in the context of self-report questionnaires, particularly in comparison to the widely used VAS. For this study we administered the English-language VNS as part of a baseline questionnaire given to 2 sets of applicants to self-management programs. All participants also were given the VAS in the same questionnaire.

MATERIALS AND METHODS

Participants. All individuals who enrolled in a self-management program for arthritis or chronic disease were given baseline questionnaires that included both the VAS and the VNS. This was a convenience sample, which took advantage of a larger separate study designed to compare an arthritis self-management program (ASMP) with a chronic disease self-management program (CDSMP). As part of the large study, questionnaires were to be administered to all potential participants at baseline (i.e., before being randomized to the CDSMP or ASMP). In 2002 through 2004, subjects were recruited from the greater San Francisco Bay Area through talks given in community locations such as seniors' centers and libraries, via public service announcements, and through disease-specific organizations such as The Arthritis Foundation. The number of participants in the initial study of the English VNS was 175. Subsequently an additional 192 participants were included in a slightly modified replication study, where the order of the VNS and VAS pain scales was reversed. A total of 418 questionnaires were mailed to potential participants in the larger self-management study, of which 367 (88%) were completed and returned.

Questionnaires and variables. The VNS was already to be included in the questionnaires for the larger self-management study, and we consequently added the VAS (separated from the VNS by several pages). This allowed us to compare the 2 scales under the same conditions as they are likely to be used by ourselves and other researchers. The VNS prompt was, "Please circle the number below that describes your pain in the past 2 weeks." For the VAS, we asked, "Please mark an "X" on the line below that describes your pain in the past 2 weeks." Among the other variables in the questionnaires, there was a 5-point measure of self-reported overall general health (excellent to poor) and

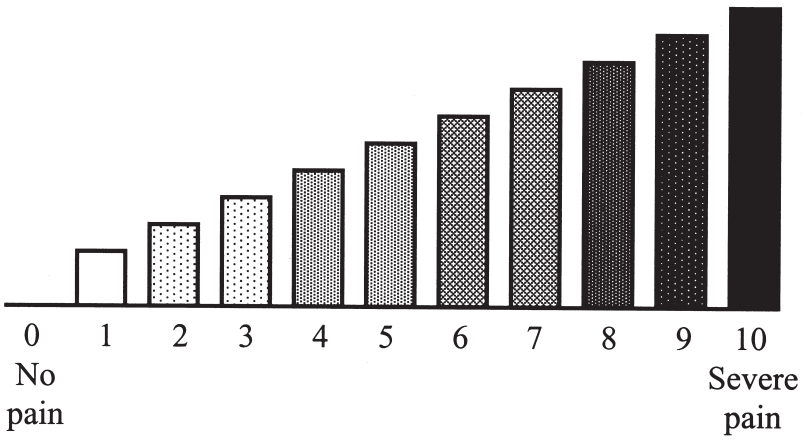

Figure 1. The visual numeric scale (VNS).

a 4-item Health Distress Scale modified from the 6-item Medical Outcomes Study (MOS) Health Distress Scale. These measures have been validated as described[17,18]. The design of the original study precluded us randomizing subjects to receive either the VAS or VNS. Because the VNS was to be used as an outcome variable for the larger study, it had to be included for all participants. The first 175 participants each received a 6-page questionnaire that included the VNS, followed by the VAS 3 pages later. In the replication study, an additional 192 participants were given a baseline questionnaire where the order of the 2 pain measures was reversed, so that the VAS was followed by the VNS. Questionnaires were sent and returned in the mail, with only written instructions on how to complete the questions, once again replicating our usual methods of administering self-report questionnaires. As part of the larger study, participants were also mailed followup questionnaires 6 months after completing the program. These followup questionnaires included the VNS, although the VAS was not included in order not to place unnecessary burdens on the participants.

The main purpose of the replication study was to test if the order of the VAS and VNS within the questionnaire affected the reporting, so not all analyses in the initial study were repeated for the replication study.

For a test-retest study, 42 participants in a separate arthritis self-management program completed the VNS twice, along with numerous other variables, one to 7 days apart.

Data analyses. Descriptive univariate statistics were computed for the VNS, as well as the VAS. The distributions were plotted for both pain measures and the Shapiro-Wilk W statistic was used to test for normality. The rate of skipping or providing invalid responses was examined. These analyses were repeated for the replication sample of participants to determine if the order of presentation of the 2 pain scales had any effect on the results, and t tests were used to compare the mean completion rates.

Prior studies of the VAS had found a uniform rather than a normal distribution, so Spearman as well as Pearson correlations were used to compare the VNS to the VAS. We hypothesized correlations between the VNS and VAS would be at least 0.7, which would indicate relatively high construct validity[19].

In the initial sample, coding for the VAS and the VNS was first done by a trained research assistant. One of the senior investigators subsequently also coded the 2 pain measures from the same questionnaires, and the results were compared. Huskisson has noted that the 21-point scale is probably the maximal level of discrimination people can make in differentiating levels of pain[20]. Consequently, we considered deviations in measurements greater than 0.5 cm on the VAS to be a coding error, and any smaller measurement error to be inconsequential. For the VNS, a coding error would be the entry of a number different than the one circled. For both the VAS and VNS, the entry of a blank when there was a valid response or entry of a number when the scale had been skipped would be considered coding errors.

Within the initial sample, both the VAS and VNS were compared to the health distress and overall general health measure using both Pearson and Spearman correlations. These 2 variables were included as outcomes in a study of the Arthritis Self-Management Program (ASMP), and we would expect them to be correlated with pain in patients with arthritis. Sherbourne found correlations of 0.58 and 0.54 between current health and health distress, respectively, and the MOS severity scale[19], while we found 0.41 and 0.36 correlation between a 20-point numeric pain scale and current health and health distress, respectively[17]. Based on those findings we hypothesized that the correlations between the VNS and health distress and overall general health would be moderate correlations of at least $r = 0.40$.

As part of the larger study of the efficacy of ASMP, we obtained data on the VNS 4 months after the intervention. For those who had been randomized to the ASMP, we looked at changes in the VNS to test its sensitivity to change. T tests were used to compare baseline and 4-month scores. In other studies we have found effect size changes (mean change divided by the baseline standard deviation) of ≥ 0.25 to be of significance to patients[21]. Thus we hypothesized that we would obtain statistically significant ($p < 0.05$) changes of that magnitude, and that such changes would confirm the sensitivity of the VNS.

To examine test-retest reliability, the VNS response at Time 1 was correlated with the VNS response at Time 2 for 42 separate participants.

RESULTS

Samples. The 175 self-management program participants in the initial sample had a mean age of 64.9 years (range 19–82) and a mean education level of 15.2 years (Table 1). Seventy-eight percent of participants had joined the study primarily because of their arthritis and 28% because of other chronic diseases (e.g., cardiovascular disease, diabetes, or lung disease). Of all participants, 66% had osteoarthritis (OA) and 13% had rheumatoid arthritis (RA). They were 84% non-Hispanic Caucasian and 84% were female; 50% were married, 16% were widowed, and 35% were single or divorced.

The 192 participants in the replication sample had a mean age of 62.8 years (range 22–83) and a mean education level of 15.2 years. They were 88% non-Hispanic Caucasian and 86% were female; 48% were married, 23% widowed, and 28% single or divorced. Seventy-one percent had OA and 14% had RA. Thus both populations can be described as relatively well educated, predominantly Caucasian and female older populations with a high frequency of arthritis. There were no significant differences between the 2 samples, although the percentage of divorced and the percentage with OA showed a trend toward significant differences ($p = 0.054$ and $p = 0.089$, respectively), with the replication sample being higher in both cases.

Descriptive statistics. In the initial sample of 175 (with the VAS first), the Visual Numeric Scale had a mean of 5.6 (SD 2.5, N = 174). Both the possible and observed ranges were 0 to 10. When measured to the nearest millimeter, the VAS had a mean of 51.3 (SD 25.9, N = 157), with observed range of 0 to 98. When rounded to the nearest 10 and converted to a 10-point scale, the mean VAS was 5.2 (SD 2.6, range 0–10).

For the 192 participants in the replication sample (with the VNS first), the mean VNS was 5.4 (SD 2.5, range 0–10, N = 192). The mean VAS was 52.5 (SD 27.5, range 0–100, N = 177).

Figures 2 and 3 show the distributions of the VAS and VNS, respectively. The VAS is rounded to the nearest 10 mm,

Table 1. Sample characteristics and descriptive statistics.

Variable	Initial Sample, N = 175	Replication Sample, N = 192
Age, yrs (SD, range)	64.9 (13.0, 19–82)	62.8 (12.3, 22–83)
Female, %	84	86
Non-Hispanic Caucasian, %	84	88
Married, %	50	48
Single or divorced, %	35	28
Widowed, %	16	23
Mean years of education (SD, range)	15.2 (2.7, 9–22)	15.2 (2.7, 6–22)
Visual numeric pain (VNS) mean (SD, range)	5.6 (2.5, 0–10)	5.4 (2.5, 0–10)
Visual analog pain (VAS) mean (SD, range)	51.3 (25.9, 0–98)	52.5 (27.5, 0–100)
Health distress, mean (SD, range)	2.2 (1.2, 0–5)	2.2 (1.3, 0–5)
Overall self-reported health (SD, range)	3.1 (0.95, 1–5)	3.1 (0.93, 1–5)

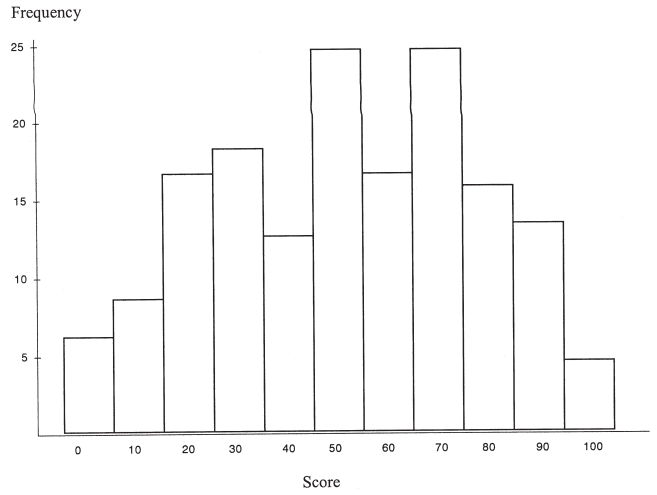

Figure 2. Distribution of responses to the visual analog scale for pain (N = 157). Scores are rounded to nearest 10.

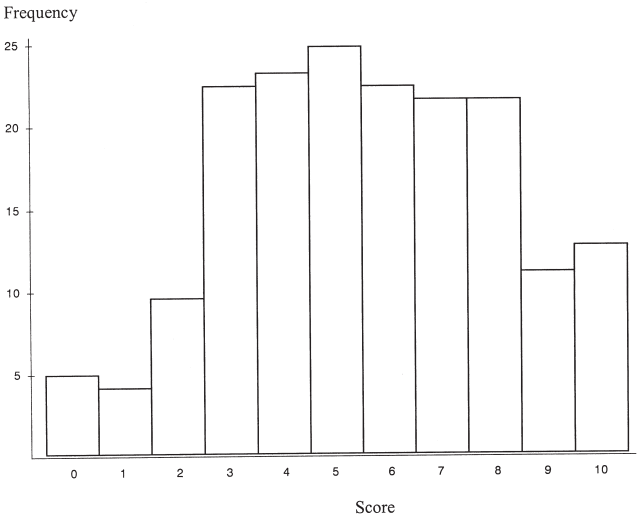

Figure 3. Distribution of responses to the visual numeric scale for pain (N = 174).

so the 2 graphs are comparable. Both scales showed deviation from the normal distribution. The Shapiro-Wilk W statistic for normality was 0.97 for VNS ($p < 0.001$) and 0.97 for VAS ($p = 0.002$). Thus, as in previous studies of the VAS[20], there was a tendency for a uniform distribution (i.e., distributed more or less equally across the range), rather than normal.

In the replication sample, both measures also deviated from normality (Shapiro-Wilk W 0.95 for VAS and 0.97 for VNS; $p < 0.001$ for both measures).

Construct validity (convergent and discriminant). In the initial sample the Pearson correlation between the VAS and VNS was $r = 0.85$ ($p < 0.001$) and Spearman correlation was 0.84 ($p < 0.001$). For the replication sample, the Spearman correlation between the 2 pain measures for this sample was 0.78 ($p < 0.001$), while the Pearson correlation was 0.76 ($p < 0.001$).

On a 6-point, 4-item health distress scale (0–5), from "health distress none of the time" to "health distress all of the time," the participants in the initial sample had a mean of 2.2 (SD 1.2). When correlated with health distress, the VAS had Spearman correlation of 0.50 and the VNS 0.62. Pearson r was 0.51 for the VAS and health distress, while the VNS had $r = 0.63$ (both $p < 0.001$).

The mean for the 5-point general health scale (1 = excellent to 5 = poor) was 3.1 (SD 0.95). Spearman correlations of the VAS and VNS to self-reported general health were 0.30 and 0.34, respectively; Pearson correlations were 0.28 and 0.35 (p for all < 0.001).

Sensitivity to change. A subset of participants was randomized to participate in an ASMP as part of a larger study. These participants received followup questionnaires, which included the VNS, 4 months after completing the course. In the initial sample, the VNS had a baseline mean of 5.27 (SD 2.81) with a mean change of –0.77 (SD 2.16) at 4 months. Thus the change in VNS was a reduction in reported pain of 0.28 effect size ($p = 0.007$) for ASMP participants (N = 61).

Test-retest reliability. The Pearson correlation coefficient (r) for the test-retest sample (N = 42) was 0.92 ($p < 0.001$). The responses were collected from the same 42 individuals one to 7 days apart (mean 2.6 days, SD 1.9).

Completion rates. Of the 175 participants given a questionnaire with the VNS followed by the VAS, 174 (99%) completed the VNS and 157 (90%) completed the VAS ($p < 0.001$ for a t test of the completion difference score). Of those not completing the VAS (18 cases), 7 respondents skipped the scale completely (left blank), 5 put two marks or a range, and 6 wrote an answer on the line, such as "mild pain" or "most of the time." The single invalid answer for the VNS consisted of a respondent marking 2 numbers and indicating a range between the 2.

In the replication sample of 192 participants, 177 (92%) completed the VAS, while all 192 (100%) completed the VNS ($p < 0.001$). Of those who did not complete the VAS (N = 15), 4 wrote a description on the line and 11 left the line blank. There was no significant difference between the 2 samples in the VAS completion rates (chi-square test).

Coding errors. When the coding performed by the research assistant was redone by the investigator and the results were compared (initial sample), 9 coding errors > 0.5 cm (5%) were discovered for the VAS, but no coding errors had been made with the VNS. This difference was statistically significant ($p = 0.001$, chi-square test). The coder also reported that the measurement necessary to code the VAS took about 6 times longer than coding the VNS (roughly 2 per minute vs 12, not including the time to locate the questions on the questionnaire).

DISCUSSION

Summary of results. When completed by patients with arthritis and chronic disease enrolled in self-management programs,

the English-language VNS for pain had means and distributions similar to those of the previously described Spanish-language VNS[14]. There were also fewer invalid responses for the VNS from the relatively highly educated English-language sample in comparison to the Spanish sample. Means and distributions were also similar to the VAS frequently used for pain when the 2 scales were included in the same questionnaire; whereas invalid answers were more frequent for the VAS regardless of the order in which the 2 scales were presented. Coding also proved to be more accurate for the VNS compared to the VAS. Correlations between the VNS and VAS were quite high, while correlations to the independent but related measure of overall health and health distress were similar and moderately high. The VNS also was sensitive enough to show statistically significant reductions in pain among arthritis patients several months after completing an arthritis self-management course.

The completion rate (99.7% across the 2 samples) was slightly higher than for the Spanish-language VNS (94%), which may be a consequence of the high education level of the English-language participants (mean of 15.2 yrs of education compared to 8.7 for the Spanish sample[15]).

The distribution of the VNS (Figure 3) showed no tendency to cluster at preferred numbers (such as 3, 5, or 7), as Huskisson[22] suggested might occur with a pure numeric rating scale. The VNS also showed a similar distribution compared to the more common VAS, with a slightly lower mean when the VAS was rounded to a 0 to 10 scale. Because of the deviation from normality, especially for the VAS, Spearman rank-order correlations are preferable. However, the very small differences between the Pearson and Spearman correlation coefficients suggest that either could be validly used for these 2 measures of pain. The Spearman correlations of 0.84 and 0.78 between the VNS and VAS, in the initial and replication samples, respectively, suggest good construct validity[19]. The high test-retest correlation for the VNS (r = 0.92) indicates strong test-retest reliability.

Based on the missing, skipped, and invalid responses, the results tend to confirm Huskisson's finding that the VAS may be confusing to some people. There was no such tendency with the VNS. The higher likelihood to skip or give invalid answers to the VAS compared to the VNS remained true regardless of the order in which the 2 measures were given. On this basis, in questionnaire or Internet studies where patients or subjects may need to answer the questions without supervision, the VNS has a strong advantage over the VAS.

In addition to being more likely to be skipped by participants, the VAS was more likely to introduce coding errors. Such errors could be reduced by a Web-based VAS; when Jamison, *et al*[24] implemented an electronic VAS the results were similar to the traditional pencil and paper VAS. However, our experience has been that the VNS is much easier to implement for Web-based use.

The correlations between both the VNS and the self-reported general health item were moderate (Spearman r = 0.35) and significant, but did not meet our criterion of r = 0.4. The correlation between the VNS and health distress was stronger (Spearman r = 0.62) and did meet our criteria. As we might expect in a population of patients with arthritis, this is where health distress may, in part, be a result of the pain from arthritis. In contrast, pain may be a less important factor contributing to self-reported overall health. Studies that compared pain to health distress and overall general health consisted of subjects with chronic disease[17] or with 4 non-arthritis "tracer" conditions (diabetes, hypertension, heart disease, and depression)[19], so pain may have been less of a factor in contributing to health distress in those studies. The slightly better but similar correlations of the VNS with self-reported general health and with health distress, as compared to the VAS with those 2 variables, suggest that the VNS has validity at least equal to that of the VAS. In general, the high correlations between VNS and VAS, as well as the moderately high correlations between the VNS and health distress, among patients with arthritis, suggest good construct validity.

A number of researchers have looked at the sensitivity of the VAS. Sensitivity can be measured by applying pain to individuals (e.g., heat or pressure), then examining changes in VAS responses[23]. Another approach is to administer analgesics and placebo to test how well the scales measure improvement in those receiving the anti-pain drugs[25,26]. In general the VAS has been found to be sensitive. It was not possible to measure sensitivity of the VNS directly using these methods. Nor would the "real-world" design of the questionnaires permit us to include additional pain measures. However, following a self-management intervention, the VNS did show significant differences in pain (with an effect size found to be relevant to patients in another study[20]). Along with the correlations between the VNS and VAS (the latter already having been shown to be sensitive), this finding argues for the sensitivity of the VNS.

Limitations. It is possible that having both the VNS and VAS in the same questionnaire, even though separated by several pages, may have influenced the responses to the second measure, thus slightly inflating the correlations between the 2 measures. We concede this possibility, but believe that the 2 measures are sufficiently different (circling a number vs marking a blank line) that the effect would be limited. The relative values of the VAS compared to VNS remained nearly the same despite the order of the questions. Dixon and Bird[27] found there were problems when volunteers attempted to duplicate marks on a VAS immediately after viewing a marked VAS. Thus we believe that the danger of the VAS or VNS strongly changing the responses to the other measure was minimal.

Since studies have suggested that a numeric rating scale has advantages over the VAS[6], one can legitimately ask why we did not compare the VNS to a numeric rating scale. Such a study would be useful, but there were several reasons for

comparing the VNS to VAS. We had used an English-language VAS fairly successfully in earlier studies, and its psychometric properties had been included in a book on outcome measures for patient education[17]. Our study was also intended to be a followup to a publication comparing the Spanish-language VNS to a Spanish VAS. And the similarities between the VNS and numeric rating scales would have demanded a different and more extensive research design to avoid the problem of the 2 scales strongly influencing each other. The VNS is a kind of numeric rating scale, with the same or similar response categories, and it would be much easier for participants to compare the number they circled for one before choosing a number for the other. Thus a randomized design, some participants receiving the VNS and some a numeric rating scale, would have been required to avoid this problem. A future comparison of the VNS to a numeric rating scale would prove informative, but was beyond the scope of the current study.

We used an educated sample of mainly non-Hispanic Caucasians. There were not enough low education or minority participants to determine the effect of ethnicity or socioeconomic status on the ability to complete the VNS. However, the slightly better completion rate compared to the Spanish VNS in an earlier study suggests that socioeconomic status or ethnicity may be a factor. While the select, homogeneous sample is a major limitation of this study, the similar sociometric findings for the Spanish-language VNS[15,16] suggest that the measure may also be applicable to a broader population. It would be desirable to study the English-language VNS among a more diverse sample, and in particular, see how successful the VNS is among non-native English speakers.

The distribution of the VNS is undoubtedly related to characteristics of our samples — specifically, an older population with arthritis or with other chronic diseases. A sample of the population at large would be expected to have a lower mean pain level, and the VNS might show strong skewing toward the zero end of the scale. Thus we cannot be certain that the VNS would be as useful in populations with low mean levels of pain.

Additional applications. The VNS may be modified by changing the prompts and anchor terms. For example, we asked about pain in the last 2 weeks — an appropriate measure for evaluating outcomes of arthritis programs. But the same scale could be used to measure pain "right now" or most intense pain over various time periods. The scale could also be used to assess various aspects of pain in addition to intensity (amount of distress caused by pain, how intolerable, how annoying, how much better or worse than some previous pain, sharpness or dullness, etc.). In addition to pain, VAS have also been used to measure anxiety[28-30], fatigue[31], depression[30], and self-esteem[32], among other variables. Similarly, VNS can also be used to measure other health conditions, and shortness of breath and fatigue VNS are being used in Spanish and English self-management program evaluations. Further studies of these modified VNS would be desirable to determine their properties.

The English-language Visual Numeric Scale for pain, compared to the Visual Analog Scale, appears to be easier for subjects to understand and answer, easier to administer with large numbers of subjects using paper questionnaires, easier to code, and less prone to coding errors. Distributions and correlations between the VAS and VNS and other related variables suggest that the VNS is at least as good as the often studied VAS in measuring the underlying subjective pain variable. The VNS also showed appropriate sensitivity to changes in pain, and it appears to be an appropriate measure for use in evaluations of arthritis and other group interventions.

ACKNOWLEDGMENT

We thank the late Byron Brown and Christina Lum for advice and assistance.

REFERENCES

1. Jensen MP, Karoly P. Self-report scales and procedures for measuring pain in adults. In: Turk D, Melzack R, editors. Handbook of pain assessment. New York: Guilford Press; 2001:15-34.
2. Gracely RH, McGrath F, Dubner R. Ratio scales of sensory and affective verbal pain descriptors. Pain 1978;5:5-18.
3. Scott J, Huskisson EC. Graphic representation of pain. Pain 1976;2:175-84.
4. Ferraz MB, Quaresma MR, Aquino LR, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. J Rheumatol 1990;17:1022-4.
5. Price DD, Harkins SW. Combined use of experimental pain and visual analogue scales in providing standardized measurement of clinical pain. Clin J Pain 1980;3:1-8.
6. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. Ann Rheum Dis 1978;37:378-81.
7. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain 1986;27:117-26.
8. Wong DL, Baker CM. Pain in children: comparison of assessment scales. Pediatr Nurs 1988;14:9-17.
9. Freyd M. The graphic rating scale. J Educ Psychol 1923;14:83-102.
10. Testler MD, Savedra MC, Holzemer WL, Wilkie DJ, Ward JA, Paul SM. The word-graphic rating scale as a measure of children's and adolescents' pain intensity. Res Nurs Health 1991;14:361-71.
11. McCaffery M, Beebe A. Pain: clinical manual for nursing practice. St. Louis, MO: Mosby; 1989.
12. McGrath PA, Seifert CE, Speechley KN, Booth JC, Stitt L, Gibson MC. A new analogue scale for assessing children's pain: an initial validation study. Pain 1996;64:435-43.
13. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. Pain 1999;83:157-62.
14. Champion GD, Goodenough B, von Baeyer CL, Thomas W. Measurement of pain by self-report. In: Ginley GA, McGrath PJ, editors. Measurement of pain in infants and children. Progress in pain research and management, vol. 10. Seattle: IASP Press; 1998:123-60.
15. Lorig K, González VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. Med Care 1999;37:957-63.
16. Gonzalez VM, Stewart A, Ritter PL, Lorig K. Translation and validation of arthritis outcome measures into Spanish. Arthritis

Rheum 1995;38:1429-46.
17. Lorig K, Stewart A, Ritter P, González V, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks, CA: Sage Publications; 1996.
18. Stewart AL, Ware JE, Sherbourne CD, Wells K. Psychological distress/well-being and cognitive functioning measures. In: Stewart AL, Ware JE, editors. Measuring functioning and well-being: the Medical Outcomes Study approach. Durham, NC: Duke University Press; 1992:102-42.
19. Sherbourne CD. Pain measures. In: Stewart AL, Ware JE, editors. Measuring functioning and well-being: the Medical Outcomes Study approach. Durham, NC: Duke University Press; 1992:220-34.
20. Huskisson EC. Measurement of pain. Lancet 1974;2:1127-31.
21. Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. JAMA 1999;282:1157-62.
22. Huskisson EC. Measurement of pain. J Rheumatol 1982;9:768-9.
23. Maxwell C. Sensitivity and accuracy of the visual analogue scale: a psycho-physical classroom experiment. Br J Clin Pharmacol 1978;6:15-24.
24. Jamison RN, Gracely RH, Raymond SA, et al. Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers. Pain 2002;99:341-7.
25. Huskisson EC, Shenfield GM, Taylor RT, Hart FD. A new look at ibuprofen. Rheumatol Phys Med 1970;10 Suppl 10:88-92.
26. Joyce CRB, Zutshi DW, Hrubes V, Mason RM. Comparison of fixed interval and visual analogue scales for rating chronic pain. Eur J Clin Pharmacol 1975;8:415-20.
27. Dixon JS, Bird HA. Reproducibility along a 10 cm vertical visual analogue scale. Ann Rheum Dis 1981;40:87-9.
28. Hornblow AR, Kidson MA. The visual analogue scale for anxiety: a validation study. Aust NZ J Psychiatry 1976;10:339-41.
29. Bond A, Shine P, Bruce M. Validation of visual analogue scales in anxiety. Int J Methods Psych Res 1995;5:1-9.
30. Tamiya N, Arake S, Ohi G, et al. Assessment of pain, depression, and anxiety by visual analogue scale in Japanese women with rheumatoid arthritis. Scand J Caring Sci 2002;16:137-41.
31. Brunier G, Graydon J. A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: visual analogue vs Likert scale. Int J Nurs Stud 1996;33:338-48.
32. Brumfitt SM, Sheeran P. The development and validation of the visual analogue self-esteem scale (VASES). Br J Clin Psychol 1999;38:387-400.