# Does the Presence or Location of Graphic Markers Affect Untrained Listeners' Ratings of Severity of Dysphonia?

***Kathy F. Nagle, †Leah B. Helou, ‡Nancy P. Solomon, and *Tanya L. Eadie,** **Seattle, Washington, †Pittsburgh, Pennsylvania, and ‡Bethesda, Maryland*

**Summary: Objective.** To determine the effect of presence and location of severity labels for different types of visual analog scales (VAS) on overall severity (OS) ratings in dysphonic speech.
**Study Design.** Experimental, between group comparisons.
**Methods.** Dysphonic and normal voice samples from male and female speakers were presented to inexperienced listeners for judgments of OS. To rate samples, listeners used an undifferentiated 100-mm VAS labeled at the extremes, a VAS with nonlinearly distributed labels as in the "beta" version of the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), or a VAS with symmetrically distributed labels as in the "official" version of the CAPE-V.
**Results.** Overall, mean OS ratings did not differ significantly across scale types, although ratings using the nonlinearly marked VAS were generally lower than those from other scales. This effect was significant for female speakers whose samples tended toward moderate OS. The ratings distribution, when compiled into 10-mm bins, differed significantly by scale type, with users of the nonlinearly marked scales skewing their ratings toward normal.
**Conclusions.** The presence and placement of labels on VAS did not significantly affect OS ratings overall, but values were significantly lower when rating female voices with the nonlinearly labeled VAS. Results indicate that professionals should specify the scale type used for rating OS and use scales consistently when comparing voices.
**Key Words:** Visual analog scale–Severity–Dysphonia–Perceptual rating.

## INTRODUCTION

Auditory-perceptual measures are part of the current "gold standard" for the clinical evaluation of disordered voice.[1] In an effort to encourage a standardized and reliable approach to evaluating auditory-perceptual features of voice, a group of voice and psychometric experts developed the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V).[2] Talkers produce two prolonged vowels, a prescribed set of six sentences, and brief conversational sample. Listeners rate the voices along a series of 100-mm lines (for easy conversion to percentages) for the following parameters: Severity, Roughness, Breathiness, Strain, Pitch, and Loudness, with an option of adding other pertinent dimensions (eg, asthenia). Raters mark a location on the line to represent the extent of perceived abnormality for each parameter. In addition, severity labels are printed beneath each line to indicate general regions for mild (MI), moderate (MO), and severe (SE). The line is intended to be a visual analog scale (VAS), and the visual labels are intended to provide an associated ordinal scale and effectively partition the line. Because of this blending of scale types, Awan and Lawson[3] referred to the scales used on the CAPE-V as "hybrid" in nature. An alternative term for a scale such as this, which uses an otherwise undifferentiated line with visual labels or markers that delineate specific areas, is a graphic rating scale.[4]

The CAPE-V form and information about its development and implementation were published on line in 2002 by the American Speech-Language-Hearing Association's Special Interest Division on Voice and Voice Disorders.[5] This original version of the CAPE-V has been used and reprinted widely in the literature.[6–8] It also has been implemented in studies addressing its reliability and validity.[9,10] In 2009, Kempster et al[2] published an updated and peer-reviewed article that also includes a reprint of the CAPE-V form. The authors note that "the form and protocol included in this article … have been modified slightly from the initial version" (p. 128). Examination of the form reveals the most obvious difference—in the original version, the visual labels for severity are placed nonlinearly with a positive skew (at approximately 10%, 35%, and 72% of the line), whereas in the revised version, the labels are centered with the two extreme labels displaced toward, but not at, the ends of each line (at 10%, 50%, and 90% of the line). Whether these discrepancies are consequential is unclear. Comparison of ratings from the original CAPE-V to a different voice evaluation scale (GRBAS[11]) revealed the nonlinear nature of the overall perception of voice quality.[9] The revised version of the CAPE-V presumably is intended to replace the original version, henceforth referred to as the "beta" version (G. Kempster, personal communication, November 10, 2010). Despite the symmetrically placed labels on the revised CAPE-V form, henceforth referred to as the "official" version, Kempster et al stated that the "ordinal ratings of *mild*, *moderate*, and *severe*, printed below the measurement line, … are positioned in a nonequidistant fashion, based

Accepted for publication December 16, 2013.
From the *Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington; †Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, Pennsylvania; and the ‡Walter Reed National Military Medical Center, Bethesda, Maryland.
Address correspondence and reprint requests to Kathy F. Nagle, Department of Speech and Hearing Sciences, University of Washington, 1417 NE 42nd St., Seattle, WA 98105. E-mail: kfnagle@gmail.com
Journal of Voice, Vol. 28, No. 4, pp. 469-475
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2013.12.011

on [a psychometrics expert's] recommendations …" (p. 127).[2] In other words, the description of the revised, official version of the CAPE-V is not consistent with the form itself.

The psychometric properties of rating scales may be compromised by combining various features of each type. As indicated by Awan and Lawson,[3] the psychophysical and statistical advantages of using a ratio scale (like VAS) may be lost when visual markers are added to it. Labeling a VAS with visual labels or markers may help guide listeners when making their judgments, but it may also introduce a bias. For example, Scott and Huskisson[4] examined differences among the uses of different types of VAS and graphic rating scales for patients rating self-perceived pain. In their multipart study, the authors compared equally distributed severity labels on a VAS to a traditional VAS marked only at the end points. Patients tended to cluster their responses around the severity labels when labels were available but uniformly distributed their pain ratings across the unlabeled VAS. Thus, the presence of the labels biased the ratings and ultimately reduced the sensitivity of these scales for assessing pain.

Although the influence of visual labels has been shown in pain ratings, it is unclear whether they affect ratings of disordered voice. Furthermore, the actual locations of visual labels along the VAS may influence ratings. Specifically, nonlinear placement of visual labels, as in the CAPE-V's *beta* version, is thought to reflect the distribution of psychophysical sensitivity to specific voice quality parameters. Alternately, symmetrical placement of severity labels on the CAPE-V's official version is thought to encourage ratings that encompass the entire scale. Given that the goal of the CAPE-V is to provide a standard evaluation form and procedure for assessing dysphonia, the availability of alternate forms presents a problem for generalizing results across clinics, clinicians, and clients.

The purpose of this study was to test the effect of the presence and locations of severity labels on ratings of normal and disordered voices by inexperienced listeners. This first line of inquiry included inexperienced listeners because they are hypothesized to be more susceptible to bias from the labels[4] and because experience affects reliability of voice ratings.[7,12] In addition, this initial study included ratings for the gestalt perception of overall severity (OS) rather than particular aspects of voice quality because it has strong face validity among individuals with voice disorders[12] and has been shown to exhibit the strongest interrater reliability of voice quality dimensions typically judged by listeners.[7,10,13,14] Specifically, this study addressed whether inexperienced listeners' ratings of OS of dysphonia differ when using an unmarked VAS, a VAS with nonlinearly placed labels, or a VAS with symmetrically placed labels.

## METHODS

Procedures were approved by the University of Washington Human Subjects Committee. All participants provided informed consent and were paid for their participation.

### Stimuli and preparation

Speech samples from 25 native English-speaking adults were selected from a database of samples recorded under identical conditions for clinical research. Demographic information including speaker sex, age, and diagnosis is provided in Table 1. Twenty-one speakers (10 males) included those

**TABLE 1.**
**Demographics of Speakers With Dysphonia and Controls (N = 25)**

| Speaker | Sex | Age | Diagnosis |
|---|---|---|---|
| Dysphonic speakers | | | |
| 1 | M | 78 | Bowing BL VFP |
| 2 | M | 64 | UL VFP, laryngopharyngeal reflux |
| 3 | M | 25 | UL VFP, possible dislocated vocal process, lesion of unknown etiology |
| 4 | M | 57 | RT VFP s/p parathyroid surgery |
| 5 | M | 53 | Papilloma, possible Parkinson disease |
| 6 | M | 64 | LT VFP s/p mediastinoscopy and upper lobectomy for lung cancer |
| 7 | M | 27 | LT VFP s/p aortic root repair and valve replacement |
| 8 | M | 70 | Laryngeal cancer, s/p partial laryngectomy and radiation |
| 9 | M | 33 | BL VF edema secondary to reflux |
| 10 | M | 60 | Clear larynx, s/p Zenker diverticulectomy |
| 11 | F | 85 | BL VFP, receiving collagen injections |
| 12 | F | 45 | RT VF cyst, LT Reinke edema |
| 13 | F | 47 | VF cyst, chronic cough for 8 years |
| 14 | F | 81 | UL VFP, muscle tension dysphonia |
| 15 | F | 63 | BL Reinke edema |
| 16 | F | 60 | VF cyst |
| 17 | F | 31 | VF nodules, subglottic stenosis |
| 18 | F | 44 | VF cyst, VF edema |
| 19 | F | 56 | Clear larynx, s/p resolved VF nodules |
| 20 | F | 38 | UL VF pedunculated polyp |
| 21 | F | 84 | VF cyst |
| Control speakers | | | |
| 1 | M | 53 | Normal voice |
| 2 | M | 37 | Normal voice |
| 3 | F | 35 | Normal voice |
| 4 | F | 53 | Normal voice |

*Notes:* Includes speaker sex, age (y), and diagnosis.
*Abbreviations:* BL, bilateral; LT, left; RT, right; s/p, status post; UL, unilateral; VF vocal fold; VFP, vocal fold paresis.