From dysphonia to dysphoria: Mokken scaling shows a strong, reliable hierarchy of voice symptoms in the Voice Symptom Scale questionnaire

From dysphonia to dysphoria: Mokken scaling shows a strong, reliable hierarchy of voice symptoms in the VoiSS questionnaire

Ian J. Deary[a,*], Janet A. Wilson[b], Paul N. Carding[b], Kenneth MacKenzie[c], Roger Watson[d]

[a]Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh, UK

[b]Department of Otolaryngology Head and Neck Surgery, Newcastle University, Newcastle-upon-Tyne, UK

[c]Department of Otolaryngology Head and Neck Surgery, Royal Infirmary, Glasgow, UK

[d]School of Nursing and Midwifery, University of Sheffield, UK

*Corresponding author: Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. Tel: +44 131 650 3452; fax +44 131 651 1771.

E-mail address: i.deary@ed.ac.uk (I. Deary).

Abstract

Symptoms of hoarseness (dysphonia) are common and often associated with psychological distress. The Voice Symptom Scale (VoiSS) is a 30-item self-completed questionnaire concerning voice and throat symptoms. Psychometric and clinical studies on the VoiSS show that it has good reliability and validity, and a clear factorial structure. The present paper presents a further advance in voice measurement from the patient's point of view. To date, there has been no examination of whether voice-related symptoms form a hierarchy; that is, whether people who suffer voice problems progress through a reliable set of problems from mild to severe. To address this question, the technique of Mokken scaling was applied to the VoiSS in 480 patients with dysphonia. A strong and reliable Mokken scale—a symptom hierarchy—was found, which included 17 of the 30 items. This new information on dysphonia shows that voice symptoms progress from voice-oriented difficulties, through practical problems, to disturbances of social relationships and mood (dysphoria). The results add information about the structured phenomenology of voice problems, further establish the relationship between voice impairment and psychosocial impairment, and suggest practical applications in the assessment of dysphonic voices.

Introduction

High vocal demand occupations, noisy social environments and continued use and abuse of tobacco and alcohol all contribute to the high prevalence of adult voice disorders in western cultures [1]. The precise number of annual voice referrals to specialists in the UK is not known. Given, however, the fact that a clear majority of voice disorders is managed conservatively by speech and language therapists [2] and that over 20,600 laryngeal examinations and operations are carried out in the UK per annum (www.hesonline.nhs.uk), well over 50,000 adults are estimated annually to have voice disorders requiring specialist attention [3]. It is hard to define accurately the world wide epidemiology of voice disorders—but they are known to comprise, for example, the commonest occupational disorder in Poland [4] and to affect, in later years, between one in five and one in three elderly people [5,6]. The impact of voice disorder is severe and wide ranging. Voice problems affect people's ability to perform in the workplace, both for professionals whose work primarily involves vocal communication and for those simply voicing for activities of daily living. They may be associated with pain, fatigue, altered self image, and distorted relationships. They may intensify with increasing age, where the concomitant reduction in hearing acuity of listeners compounds an already very frustrating state of affairs. The gold standard assessment of the severity of a voice disorder is thus typically multidimensional, incorporating evidence from an expert rater, and/or a computerised sound analysis system. The use of a patient reported outcome measure has two main advantages. First, it offers the opportunity to inform quantitatively on the wider psychosocial implications of the problem, which cannot be judged by mere

rating of voice quality. Secondly it offers a reflection on the sustained impact over a longer timescale than the ‘single snapshot’ impression of the clinic visit observations.

The VoiSS—a patient reported outcome measure—was designed to be applicable across the range of heterogeneous voice symptoms. It was developed in the UK after rigorous psychometric evaluation of its content validity, internal consistency, sensitivity to change, and psychometric structure on a series of large samples of voice patients [7,8,9]. It has three subscales, impairment (fifteen items); emotional response (eight items) and physical symptoms (seven items).

In the present study, the statistical technique of Mokken scaling is applied to the VoiSS for the first time, to offer new insights into the relationship between the individual items within the scale. Many readers of the journal are likely to be unfamiliar with the Mokken scaling procedure. Therefore, before discussing the rational for the current application, we outline the method for the non-expert. Mokken scaling was developed by Molenaar and Sijtsma [10]. The MSP programme, which searches polychotomous item banks for reliable, hierarchical scales, enables the identification of monotone homogeneity and double monotonicity among those items [11]. Monotone homogeneity is the degree to which an item’s score value reflects the underlying trait value, i.e. as the underlying trait value increases, the item score increases. Double monotonicity is the situation where item’s level of difficulty differs from another item and the slopes describing their monotone homogeneity do not intersect. For a good Mokken scale both monotone homogeneity and double monotonicity of items should be achieved. Initially, scales are

identified on the basis of the scalability of sets of items using Loevinger's coefficient of scalability (H) which should exceed 0.3 for a potential Mokken scale to be present; H ≥ 0.4 indicates a medium scale [10]. Loevinger's coefficient reflects how closely items' difficulty rates conform relative to one other. MSP also generates a diagnostic value '*Crit'* which calculates a single value from the combined H coefficients of the items retained in the analysis. *Crit* values of zero are considered to indicate perfectly non-intersecting items and values of *Crit* < 40 are considered to be the result of sampling error whereas values of *Crit* > 80 are considered to indicate violations of monotone homogeneity and double montonicity. Therefore, it is considered acceptable to include items with *Crit* values ≥ 0 or < 40 [10]. A further diagnostic, the P(++) matrix, shows the probability of obtaining items at certain points in the scale, and can be visually inspected. For an acceptable Mokken scale, the P(++) matrix should increase from right to left and from top to bottom [12]. A statistical procedure measuring reliability, analogous to Cronbach's alpha [13] in that it is used to test reliability of the scales obtained by the MSP, generates a value—Rho—which should be ≥ 0.7 for a reliable scale. A Bonferroni-type method of correction [10] is used within the MSP to estimate the probability of obtaining any scale generated by accounting for the multiple steps involved in this iterative programme. Summary scale statistics can also be generated (mean, skewness and kurtosis) to show how closely scores obtained using the final scale are normally distributed.

To date, the VoiSS has been developed using principal components analysis which, while useful for extracting unidimensional scales from multivariate datasets, merely provides

factors that are comprised of sets of items that correlate highly with one another [14]. Mokken scales have the additional feature in that a Mokken scale orders items systematically [14]. In addition, Mokken scales demonstrate, for the measurement of a latent trait, how the performance of an item, in terms of its score, varies with changes in the latent trait being measured by the overall scale [15]. Therefore, when item scores in a Mokken scale are summed, this is a measure of the order of the latent trait being investigated [13]. It is increasingly recognised that Mokken scaling adds complementary information to self-report scales which already are in wide use and whose psychometric characteristics, hitherto, have been investigated using only factor analytic-type approaches [11,16].

The general aims of the present study are to advance the measurement of voice pathology, particularly with respect to the patient's perspective on voice difficulties, and to further establish the relation between voice impairment and psychosocial impairment. Specifically, we apply the Mokken scaling procedure to the VoiSS questionnaire responses to discover whether there is a consistent natural hierarchy of voice symptoms among people with dysphonia. One specific aim is to find out of there is a place on a reliable continuum of voice-related complaints where reports of dysphoric phenomena appear.

Method

Participants

Data were obtained from an initial sample of 496 hospital-referred, dysphonic participants (353 women, 143 men) who completed the VoiSS questionnaire and whose data contributed to two previous reports on its factor structure [7,8]. The mean (SD, range) age for the 492 subjects who provided age data was 52.0 years (17.0, 16 to 88). Of these subjects, 480 provided responses to all of the 30 items of the VoiSS questionnaire and their data are analysed here. Many dysphonia diagnoses were included, with the commonest primary diagnostic labels being (in order of frequency): functional (N = 145), vocal fold paresis (N = 63), laryngitis (N = 47), Reinke's oedema (N = 29), apparent hyperfunction of laryngeal musculature (N = 27), polyp (N = 18), ventricularis plicae (N = 17), nodules (N = 15), glottic gap (N = 13), papillomata (N = 11), and several additional diagnoses with lower frequencies. The Ns are approximate, because primary diagnostic labels were applied using a variety of terminology. We did not attempt to exclude either those with structural lesions or those without, because we provided strong evidence that the distinction is not achievable on a clinical basis or on the grounds of personality traits and psychological distress—and theoretically there may also be overlap between the two categories [17].

VoiSS

The VoiSS questionnaire, a 30-item self report instrument [8], asks about voice-related problems. Each item has a stem question and a five-choice response scale: never, occasionally, some of the time, most of the time, always. Its subscales were described in the Introduction. Although it has subscales, almost all of the 30 items load highly on the first unrotated principal component of a principal components analysis [8]. This indicates

that the 30 VoiSS items are markers of an underlying general voice problem. It is not known whether the items form a consistent, natural hierarchy of severity.

Mokken scaling

Data were entered into SPSS for Windows version 13.0 and subjects with missing data were removed. The data (N = 480) were saved as a tab-delimited file with the spreadsheet option turned off to create a file that can be read by the MSP version 5.0 for Windows software, run on an IBM compatible PC. By increasing the lowerbound H value incrementally in 0.05 steps from 0 to 0.50, the number of scales obtained, the number of items they contained, and their reliability were recorded. This preliminary analysis continued until reliable scales with sufficiently high H were obtained before further analysis of monotone homogeneity and double monotonicity were carried out.

Results

From the 30 VoiSS items, a highly reliable scale with17 items was obtained (Rho = 0.93; Table 1), after checking for monotone homogeneity and double monotonicity. Table 1 shows a scale for the total data set which increases in 'difficulty' from 'Is your voice hoarse?' (mean item endorsement = 3.26) to 'Do you feel lonely because of your voice problem?' (mean 1.38). The scale is a strong Mokken scale (H = 0.49) in addition to its being highly reliable. Moving up the items from least (items with the highest mean endorsement) to most difficult (items with the lowest mean endorsement), the scale traces a process of first reporting the vocal aspects of their problem (e.g. difficulty competing against a background noise; finding the effort tiring), through the specific problems (e.g.

people not hearing you; voice ‘giving out’), to more affective aspects (e.g. feeling embarrassed, ashamed and lonely).

Discussion

The hierarchy of self-reported voice pathology runs from dysphonia to dysphoria. This Mokken scaling analysis of the voice-related problems contained in the 30-item VoiSS scale offers a clearly structured phenomenology in this large sample of heterogeneous dysphonia patients. That is, at the mild end of the hierarchy people report the actual symptoms of the voice problem, then the practical consequences appear and, finally, at the severe end of the scale, the emotional and social consequences. The VoiSS is remarkable in having within it such a strong and easily-interpretable scale, one that incorporates such a large proportion of the scale’s items by comparison with other self-report scales [11,16]. This achieves our aims of advancing the measurement of the patient’s perspective on voice pathology, and it further establishes an association between voice impairment and psychosocial impairment.

The data do not address the additional, important question, of whether personality traits such as neuroticism or negative affectivity predispose to the dysphoria found in dysphonia; neither do they address whether the dysphoria comes as a result of the dysphonia [18]. It will also be interesting to investigate whether other disorders of communication—found, for example, after stroke and in people who stutter—have a similar hierarchy. This could establish how general such a hierarchy is among those who have problems communicating.

According to the criteria for judging Mokken scales, whereby H > 0.5 indicates a strong scale, the Mokken scale derived here from the VoiSS is a strong scale (H = 0.49) [19]. It is also very reliable (Rho = 0.93) and retains 17 of the 30 items entered into the procedure. This compares very favourably with other recently derived Mokken scales for psychological instruments whereby, for example, five of the 12 items for Neuroticism in the NEO-FFI were retained in a Mokken scale [11], and nine items of the GHQ-30 were retained in a Mokken scale [16]. This points to the value of the progressive development of the VoiSS which has produced a scale with items which are largely authentic, in terms of describing different aspects of dysphonia but with little redundancy in terms of a hierarchical scale.

The utility of demonstrating this hierarchy of voice-related symptoms in the VoiSS scale is related to the underlying item-response theory of Mokken scaling. A score on any item in a hierarchy indicates which of the items have been scored and those which have not. For example, an individual who scores that they are embarrassed by their voice problem will probably have scored on all the items below that in the hierarchy such as having a hoarse voice and finding the effort of speaking tiring; however, they might not have scored on the items indicating greater severity, such as being ashamed or lonely. The possibility exists, therefore, of producing a very rapid clinical assessment of dysphonia by asking individuals a series of dichotomous (yes/no) questions to estimate severity.

In summary, these findings complement the already-extensive data on the VoiSS scale, and extend its validity. We now know that there is a progression in how people experience and report their voice problems. That progression—hierarchy—is not idiosyncratic; it is based on a clear progression of item content. The hierarchy was found within our large, heterogeneous group of patients with many different voice pathologies. It is a strength of the study that we found such a clear hierarchy in such heterogeneous patients. It might have been easy (and relatively trivial) to find such a hierarchy in very similar patients, with similar problems. However, if, among our different diagnoses, there had been different hierarchies—or some groups with no hierarchy—this would have prevented the clear emergence of a hierarchy, such as that found here. The appearance of dysphoria-related items at one end of the Mokken scale might appear to indicate that the results apply to functional more than organic diagnoses. However, our previous research showed that there were no differences in personality traits or psychological distress between organic and functional dysphonia patients, and that both groups tended to experience greater psychological distress than expected from norms [17].

Apart from the fundamental information about the ordering of voice problems, the findings are also of practical interest. They allow the practitioner to realise that, for most people, the psychosocial sequelae of voice pathology are at the ‘far end’ of the experience. Low mood, embarrassment and loneliness—dysphoria—as a result of voice problems tend to occur after a number of other stages have been passed through. Nonetheless, for patients presenting to a secondary or tertiary level of care, such issues

may be relatively frequent and our ongoing work continues to explore the role of cognitive behaviour therapy for chronic dysphonia [20].

Acknowledgements

The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative. Funding from the BBSRC, EPSRC, ESRC and MRC is gratefully acknowledged.

References

[1] Williams N, Carding PN. Occupational Voice Loss. London: Taylor & Francis, 2005.

[2] Carding PN. Voice pathology clinics in the UK. Clin Otolaryngol 2003;28:477-8.

[3] Mackenzie K, Millar A, Wilson JA et al. Is voice therapy an effective treatment for dysphonia? BMJ 2001;323:658-62.

[4] Wilczynska U, Szeszenia-Dabrowska N, Szymczak W. Occupational diseases in Poland 2007. Med Pr 2008;59:113-22.

[5] Roy N, Stemple J, Merrill RM et al. Epidemiology of voice disorders in the elderly: preliminary findings. Laryngoscope 2007;117:628-33.

[6] Turley R, Cohen S. Impact of voice and swallowing problems in the elderly. Otolaryngology Head and Neck Surgery 2009;140:33-6.

[7] Deary IJ, Wilson JA, Carding PN, et al. VoiSS: a patient derived voice symptom scale. J Psychosom Res 2003;54:483-9.

[8] Wilson JA, Webb AL, Carding PN, et al. Comparing the Voice Symptom Scale (VoiSS) and the Voice Handicap Index: structure and content. Clin Otolaryngol 2004;29:169-74.

[9] Steen IN, Webb AL, Deary, IJ, et al. Optimising outcome assessment of voice interventions II: the sensitivity to change of self report and observer rated measures. J Laryngol Otol 2008;122:46-51.

[10] Molenaar IW, Sijtsma K. MSP5 for Windows. Groningen: iec ProGAMMA, 2000.

[11] Watson R, Deary I, Austin E. Are personality trait items reliably more or less ‘difficult’? Mokken scaling of the NEO-FFI. Pers Individual Dif 2007;43:1460-9.

[12] Niemöller K, van Schuur W. Stochastic models for unidimensional scaling: Mokken and Rasch. In: McKay D, Schofield N, Whiteley P, editors. Data Analysis and the Social Sciences (pp.120–70). London: Francis Pinter, 1983.

[13] Moorer P, Suurmeijer, TPBM, Foets M, et al. Psychometric properties of the RAND-36 among three chronic diseases (multiple sclerosis, rheumatic diseases and COPD) in the Netherlands. Qual Life Res 2001;10:637-45.

[14] DeJong A, Molenaar IW. An application of Mokken's model for stochastic, cumulative scaling in psychiatric research. J Psychiatr Res 1987;21:137-49.

[15] Meijer RR, Baneke JJ. Analyzing psychpathology items: a case for nonparametric item response theory modelling. Psychol Methods 2001;9:354-68.

[16] Watson R, Deary IJ, Shipley B. A hierarchy of distress: Mokken scaling of the GHQ-30. Psychol Med 2008;38:575-9.

[17] Millar A, Deary IJ, Wilson JA, MacKenzie K. Is an organic/functional distinction psychologically meaningful in patients with dysphonia? J Psychosom Res 1999:46;497-505.

[18] Deary IJ, Scott S, Wilson I, White A, MacKenzie K, Wilson JA. Personality and psychological distress in dysphonia. British Journal of Health Psychology 1997;2:333-341.

[19] van Schur WH. Mokken scale analysis: between the Guttman scale and parametric item response theory. Political Analysis 2003;11:139-63.

[20] Daniilidou P, Carding P, Wilson JA, l Drinnan M, Deary V. Cognitive behavioural therapy for functional dysphonia: a pilot study. Ann Otol Rhinol Laryngol 2007; 116:717-22.

Table 1

An overview of the seventeen item VoiSS Mokken Scale. Items at the bottom are those endorsed first ('easiest' items) and those at the top are those that tend to be endorsed only when those below them have been endorsed ('hardest' items).

Item No.	Mean Score	H[a]	Item stem content
44	1.38	0.54	Do you feel lonely because of your voice problem?
43	1.59	0.49	Are you ashamed of your voice problem?
31	1.69	0.47	Do people seem irritated by your voice?
27	1.78	0.47	Does you voice problem put a strain on your family and friends?
42	1.78	0.53	Does you voice make you feel incompetent?
23	2.06	0.51	Does your voice problem make you feel stressed and nervous?
1	2.09	0.45	Do you have difficulty attracting attention?
17	2.25	0.51	Do you feel miserable or depressed because of your voice problem?
21	2.29	0.53	Are you embarrassed by your voice problem?
40	2.42	0.42	Does your voice 'give out' in the middle of speaking?
8	2.47	0.51	When talking in company do people fail to hear you?
22	2.53	0.53	Do you find the effort of speaking tiring?
16	2.56	0.52	Do you have problems talking on the telephone?
15	2.77	0.45	Do you have a weak voice?
34	2.86	0.40	Does your voice feel creaky and dry?
24	2.87	0.54	Do you have difficulty competing against background noise?
7	3.26	0.43	Is your voice hoarse?

Note.
Scale coefficient H = 0.49; n = 480; p (adjusted) = 0.000099; reliability = 0.93; mean = 38.64; standard deviation = 13.58; skewness = 0.76; kurtosis = 0.19.

[a]For description of H coefficient, see text.