Development of a new Clarity of Auditory Imagery Scale

JOHAN WILLANDER AND STEPHAN BARALDI
Stockholm University, Stockholm, Sweden

In the psychological study of auditory imagery, instruments for measuring vividness or clarity have existed for some time. The present article argues that existing scales are ambiguous, in that clarity and vividness of auditory imagery are addressed simultaneously, and that empirical validations of those scales suffer from inadequate methods. The aim of the present study was to develop a new psychometric scale, the Clarity of Auditory Imagery Scale, measuring individual differences in clarity of auditory imagery. Drawing on previous literature, 16 items were generated, forming an initial item pool that was presented to 212 respondents. The hypothesized single dimensionality inherent in the data was confirmed using Velicer's (1976) minimum average partial test and parallel analysis. Also, data were factor analyzed, extracting a stable one-factor solution including all 16 items. The internal consistency of the final scale was satisfactory (coefficient alpha = .88). Other properties of the questionnaire, such as test–retest reliability, remain to be established.

Mental imagery may be investigated using various methods (see, e.g., Bensafi et al., 2003; Halpern, Zatorre, Bouffard, & Johnson, 2004; Kosslyn, Reiser, Farah, & Fliegel, 1983; Marks, 1973). One of the main methods in the investigation of the phenomenological experience of mental imagery is self-report questionnaires, such as the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973) and the Vividness of Olfactory Imagery Questionnaire (VOIQ; Gilbert, Crouch, & Kemp, 1998). Typically, the subjective experience of mental images concerns vividness (i.e., a combination of clarity and liveliness) or clarity (i.e., a combination of brightness and sharpness; see Marks, 1999; McKelvie, 1995) of mental images, a topic early highlighted in studies by Galton (1880, 1907) and Betts (1909). In the Galton protocol (the “breakfast-table questionnaire”; Galton, 1907), participants were asked to visually consider “your breakfast-table as you sat down to it this morning” (p. 58; though the breakfast table served only as an example of something familiar) and to write down what they could see in their mind's eye with respect to three dimensions—illumination, definition, and coloring.

Galton's (1907) work also included auditory imagery (e.g., the beat of rain against the window panes, the slam of a door), which was assessed in a qualitative form similar to that used for visual imagery. We argue that over a century later, an adequate scale for measuring individual differences in auditory imagery is still unavailable. To the best of our knowledge, three scales have been developed to date that seek to evaluate auditory imagery: (1) Betts's Questionnaire upon Mental Imagery (QMI), (2) the shortened

form of Betts's QMI, and (3) the Auditory Imagery Scale (AIS). However, several critical issues may be identified in these scales. The main concern with the first two is that they are conceptually ambiguous in that clarity and vividness are addressed simultaneously. In addition, two further issues may be identified in the three scales—inconsistent labeling of scale levels and inadequate criteria for establishing dimensionality. To address these issues, we developed a new scale that measures only clarity of auditory imagery by using the minimum average partial test and parallel analysis in order to establish its dimensionality: the Clarity of Auditory Imagery Scale (CAIS).

On the basis of Galton's original instrument, Betts (1909) developed the QMI. This questionnaire measures an individual's ability to form mental images as assessed by quantitative ratings of clarity and vividness. Compared with Galton's questionnaire, the QMI includes a broader spectrum of questions addressing the many different types of imagery—that is, visual, auditory, olfactory, gustatory, cutaneous, kinesthetic, and organic imagery.

The QMI includes 20 questions pertaining to the auditory domain (Betts, 1909). Individuals are asked to imagine (1) various aspects of a teacher's voice (4 questions), (2) a familiar tune played by different instruments or sung (4 questions), and (3) various sounds—for example, the whistle of a locomotive, the rattling of a newspaper, or the sound of a clock striking (12 questions). Each representation is rated on a 7-point scale (1 = *perfectly clear and as vivid as the actual experience*; 2 = *very clear and comparable in vividness to the actual experience*; 3 = *moderately clear and vivid*; 4 = *not clear or vivid but recogniz-*

J. Willander, jwill@psychology.su.se

able; 5 = *vague and dim*; 6 = *so vague and dim as to be hardly discernible*; 7 = *no image present at all, you only are knowing that you are thinking of the object*).

The main objection to the QMI is that it was designed to measure clarity and vividness of auditory imagery simultaneously. Since vividness may be defined as a combination of clarity and liveliness (see Marks, 1999; McKelvie, 1995), clarity should be addressed rather than general vividness in order to decrease potential ambiguity. Furthermore, the verbal anchors are unclear. For example, the precise difference between *not clear or vivid but recognizable* and *vague and dim* is highly unclear. Indeed, Richardson (1977) concluded that in order to obtain reliable and valid self-report scales, it is necessary to "improve the scales . . . by reducing the ambiguity that still exists in many of the items" (p. 42). With regard to item selection, no statistical analysis aimed at obtaining a proper set of items was reported. The stability of the QMI over time is also questionable. For example, White, Ashton, and Brown (1977) reported a test-retest reliability of .46 (1 year between tests) for the auditory subscale of QMI, which is substantially lower than the test-retest reliability of .74 (6 weeks between tests) reported by Evans and Kamemoto (1973).

The shortened form of Betts's QMI is a modified version of Betts's (1909) original QMI in which each subscale (i.e., modality) comprises five items (Sheehan, 1967). The selection of items was carried out by first applying principal components analysis (PCA) on data from 280 participants who had completed the original QMI. Next, five items were selected that displayed a combination of loading highly on the main component for each modality, had similar means and standard deviations, and did not indicate sex differences. As with the original QMI, the shortened form of QMI also suffers from the problem of measuring clarity and vividness of auditory imagery simultaneously. Further, the procedure used by Sheehan for establishing the scale's dimensionality is unconventional compared with modern standards (cf. O'Connor, 2000). Items were classed into modality-specific groups following the original QMI before being subjected to component analysis. Next, Sheehan selected a small sample of modality-specific items on the basis of similarities in means and standard deviations; however, the number of selected items (i.e., five per modality) appears to have been arbitrary.

The third instrument aimed at the investigation of auditory imagery is the AIS. In the AIS, which is modeled after Betts's (1909) QMI, individuals are asked to imagine seven different sounds (e.g., the sound of water dripping, footsteps coming up stairs, your favorite piece of music). All imagined sounds are rated on a 4-point scale (1 = *very clear sound/noise*; 2 = *moderately clear sound/noise*; 3 = *vague sound/noise*; and 4 = *no sound/noise at all*).

The unequal-length Spearman-Brown split-half reliability coefficient of the scale was reported to be .80, suggesting that AIS is reliable (Gissurason, 1992). However, three major problems can be identified in AIS. First, the method for establishing the dimensionality of the scale was inadequately reported, rendering the validity of AIS questionable. Second, the labeling of the verbal anchors is

again unclear. More precisely, scale levels 1 and 2 are labeled as different levels of "clear," whereas scale level 3 is labeled "vague." Indeed, Gissurason noted that the scale is preliminary and needs further validation. Third, it is unknown how the specific items were generated.

In sum, two of the three existing scales are conceptually ambiguous in that they address both clarity and vividness simultaneously. Also, the response formats used in all three scales add to their ambiguity. In addition, it is unclear how the initial item pools were generated, and the dimensionality of these scales has been established using inadequate criteria for selecting the optimal number of underlying components or factors.

In view of the previous statements, the present study aimed to develop a new questionnaire—the CAIS—that (1) was designed to measure only clarity of auditory imagery, (2) was designed to avoid the criticisms raised with previously used response scales, (3) has final items that were selected using more accurate criteria for establishing the dimensionality of the scale, and (4) reports full details of the item-selection procedure.

METHOD

Guided by existing literature and by informal discussions and interviews with peer scholars and students, we generated 16 items reflecting a range of common auditory experiences. During the construction of the item pool, an important focus was to sample sounds that are elements in people's everyday lives and that comprehensively cover a range of sounds that people often hear in Western society (Reckase, 1996). A questionnaire containing the 16 items, formulated in Swedish, was filled out by 223 students at the Department of Psychology at Stockholm University in exchange for course credit. Respondents were instructed to rate how clearly they could imagine the sounds presented in the 16 items. Each imagery representation was rated on a 5-point Likert-type response format with verbal labels on the end points (1 = *not at all*; 5 = *very clear*). In total, 212 individuals with complete data were located and included in the analyses, yielding an internal attrition rate of 5%. Fifty-eight of the respondents were men (27%), and the respondents' age ranged between 18 and 53 years ($M = 25.92$, $SD = 6.64$). Most respondents completed the 16-item questionnaire in approximately 5 min.

RESULTS

As for data normality, Items 2, 5–11, and 13–16 had skewness values ranging between -0.18 and -0.92 , whereas Items 1, 3–4, and 12 had skewness values ranging between -1.12 and -1.65 . Hence, 4 of the 16 items were slightly outside the generally recommended range of ± 1 (Bulmer, 1979). Also, the zero-order correlation matrix was inspected for bivariate correlations above .30 in order to establish factorability (Tabachnick & Fidell, 2001). Correlations ranged between .11 and .54, with 63 out of 120 correlations exceeding .30, indicating that the data were appropriate for factor analysis.

Before proceeding with factor analysis of the data, the optimal number of components to be extracted had to be decided. The number-of-factors issue has long been debated (Gorsuch, 2003), with some procedures being portrayed as highly problematic. For example, the eigenvalues-greater-than-one rule, as well as scree-plot examination (both regularly implemented in common sta-

tistical packages), are per se unreliable methods for estimating the optimal number of underlying components or factors. Fabrigar, Wegener, MacCallum, and Strahan (1999) noted that the eigenvalues-greater-than-one procedure is often misapplied to eigenvalues of the reduced correlation matrix (i.e., with communality estimates in the diagonal) rather than to eigenvalues of the unreduced correlation matrix (i.e., with unities in the diagonal). Fabrigar et al. also noted that this procedure is a mechanical one, and that the distinction between a factor with an eigenvalue of 1.01 as nontrivial and a factor with an eigenvalue of 0.99 as trivial is arbitrary. In addition, this procedure is prone to both overestimate and underestimate the number of components (Zwick & Velicer, 1986), as well as to generate unreliable components (Cliff, 1988). Indeed, Velicer, Eaton, and Fava (2000) demonstrated that the eigenvalues-greater-than-one rule is extremely inaccurate and recommended that it not be used. As for the scree test, the practice of visually inspecting scree plots for sharp demarcations between eigenvalues for nontrivial versus trivial components is in many situations unreliable (Gorsuch, 2003; Streiner, 1998).

Several alternative procedures for determining the optimal number of components have been proposed. Velicer et al. (2000) suggested that a combination of a minimum average partial (MAP) test (O'Connor, 2000; Velicer, 1976) and a parallel analysis (PA) (O'Connor, 2000) is one of the most accurate and reliable decision rules. In contrast to mechanical rules of thumb, these procedures are statistically based (O'Connor, 2000). Instead of using standardized unities of contributed variance from each variable to the component extraction as an indication of which components are or are not trivial, the MAP test looks at the systematic and unsystematic variance remaining after extractions of increasing numbers of components, and PA searches for the number of components that explain more variance than do components extracted from random data with equal numbers of variables and cases.

In line with these recommendations, we performed a MAP test, followed by a PA in order to determine the optimal number of factors to extract from the data. For both procedures, we used programs for the SPSS Matrix-End Matrix environment written by O'Connor (2000). The MAP program generates a PCA and a series of partial correlation matrices. In $k - 1$ (where k is the number of variables) steps, principal components are partialled out of the correlation matrix, and the off-diagonal average squared coefficient is computed for the resulting correlation matrices. For the respective $k - 1$ roots, average squared partial correlations and partial correlations raised to the fourth power are then lined up. The root with the lowest squared partial correlation signifies the number of components to be extracted from the data.

The PCA on the initial pool of 16 items generated four components with eigenvalues greater than 1; however, the lowest average squared partial correlation (.0163) generated by the MAP test was at the first root. This was also the case for correlations raised to the fourth power (.0006).

Hence, the MAP test clearly indicated that only one factor should be extracted from the data.

Next, we subjected the data to PA. This procedure involves computing eigenvalues from a large number (mainly restricted by CPU power limit) of random data sets that equal the observed data in terms of numbers of cases and variables. Eigenvalues are then computed for the observed data matrix (with component analysis extracting as many factors as the number of variables, and with the number of iterations fixed at zero) and are compared with eigenvalues from the random data-set matrices corresponding to the 95th percentile. Accordingly, we generated 1,000 random data sets with 212 cases and 16 variables. Factors are to be retained when the i th eigenvalue from the observed data is greater than the i th eigenvalue from the random data.

The results from the PA mirrored the results from the MAP test. The observed data eigenvalue at the first root (5.12) was greater than the mean random data eigenvalue at the 95th percentile (1.60); however, the observed data eigenvalue at the second root (0.73) was less than the mean random data eigenvalue at the 95th percentile (1.47). Hence, the PA also clearly indicated that only one factor should be extracted from the data.

Next, we analyzed the data with exploratory factor analysis (EFA). The rationale behind choosing EFA versus PCA is usually of minor importance, since, given a large sample size, they often lead to the same conclusions (Gorsuch, 2003); however, a number of arguments speak in favor of EFA. For instance, EFA is less susceptible than PCA to extracting factors from data with little or no correlation (Gorsuch, 2003). Also, when there is reason to believe that there is error in measurements—such as in the case of developing instruments for measuring psychological constructs (i.e., error-free latent variables)—EFA is more suitable, since factors reflect such latent variables (Gorsuch, 1990; Worthington & Whittaker, 2006), and the goal is then to explain common variance and not total variance (Tinsley & Tinsley, 1987). Also, given the exploratory nature of the present analysis, principal axis factoring was used for extraction.

On the basis of the results from the MAP test and the PA, only one factor was extracted. Kaiser's measure of sampling adequacy was .88, which is well above the threshold of .60 required for good factor analysis (Tabachnick & Fidell, 2001). Factor loadings for the 16 variables that were entered into the analysis ranged between .40 and .67; thus, all factor loadings were above the threshold of .32 recommended by Tabachnick and Fidell. Communality estimates for the included variables were generally low to moderate, ranging from .16 to .44, indicating low systematic variance and high error variance in the respective items. The extracted single factor explained 31.63% of the scale's total variance.

Finally, a reliability test was performed in order to estimate the final scale's internal consistency. The scale's estimated reliability was satisfactory, with a coefficient alpha of .88 (standardized), which is well above the gen-

erally recommended threshold of .70 (DeVellis, 2003). Deleting any of the 16 items did not result in increased reliability of the scale.

A simple control for demographic background was also performed. Mean values for the final scale were calculated across the sample, correlated with the respondents' age, and compared between sexes. The correlation between CAIS and age was .07 ($p > .05$). A t test comparing CAIS means between men ($M = 3.91$, $SD = 0.66$, $n = 58$) and women ($M = 3.90$, $SD = 0.57$, $n = 154$) revealed no significant difference [$t(210) = -0.05$, $p > .05$].

DISCUSSION

In the auditory imagery literature, a number of studies are based on measuring the vividness or clarity of mental auditory representations (e.g., Brett & Starker, 1977; Okada, Matsuoka, & Hatakeyama, 2005); however, the methodological concerns underlying the development of previous scales measuring such representations call for a new valid and reliable measure. In this study, we presented just such a new measure (the CAIS). Our main objective was to develop a psychometrically robust scale with multiple indicators measuring individuals' clarity of auditory imagery. On the basis of a MAP test, PA, and common factor analysis, 16 items were included in the final scale and showed satisfactory reliability. In sum, the results suggest that the CAIS consistently measures one psychological construct—clarity of auditory imagery.

Of potential concern is the rather low communalities identified in the present investigation. The communalities for the included variables were generally low to moderate, indicating low systematic variance and high error variance—that is, low reliability in some of the variables. Scholars have emphasized the importance of keeping the number of scale items low in order to avoid inflation of the scale's reliability (Carmines & Zeller, 1979; Niemi, Carmines, & McIver, 1986); however, the items in the present analysis suffered from low reliability, as illustrated by the low communalities. We interpret this finding as emphasizing the difficulty of effectively capturing individuals' clarity of imagery as a psychological construct. The low to moderate communalities would, in fact, suggest that enough items are needed to adequately sample the target construct, and the factor analysis also indicated that all items be included in the final scale. Also, the amount of variance explained by our factor solution was just above one third of the scale's total variance. There is, however, no established criterion that is based on the amount of variance explained by a factor solution (Tinsley & Tinsley, 1987), and our results are in line with factor-analytic findings in general (cf. Hinkin, 1995). Regardless, the CAIS should be submitted to further validation in order to sharpen formulations and/or selection of items, lowering the total amount of items while increasing item reliability and explained variance.

Although we have presented evidence in support of the CAIS's psychometric properties, further validation in terms of its relationships with external variables is also

recommended. Such variables would, for example, constitute neuroimaging data. Studies have shown that auditory imagery activates brain regions (e.g., the secondary auditory cortex) also involved in regular auditory perception and that the imagery-related activations are modality specific (e.g., Zatorre, Halpern, Perry, Meyer, & Evans, 1996). Also, on the basis of VVIQ scores, Marks and Isaac (1995) found that the EEG alpha activity was attenuated in the left posterior part of the brain during visual imagery for vivid versus nonvivid imaginers, suggesting that self-rated vividness scores correspond to imagery-evoked brain activity. This taps into the central question of whether the underlying construct in the present work reflects auditory imagery or some more general (i.e., not modality-specific) imagery function; thus, it would be highly valuable to relate CAIS ratings to brain activations associated with imagery in general and auditory imagery and perception more specifically.

Another important aspect that needs further attention is the test-retest reliability of the CAIS. Test-retest reliability is an important property of any scale measuring constructs that are assumed to be stable over time. Also, an issue of further interest concerns the effects of socially desirable responding on the CAIS. For example, it would be valuable to investigate the correlation between the CAIS and self-deceptive enhancement (cf. Allbutt, Ling, Heffernan, & Shafiullah, 2008).

In conclusion, the present work proposes a new scale for measuring individual differences in clarity of auditory imagery. The presented scale was shown to be both valid and reliable by the tests conducted on it thus far, although future research for further validation and development is called for.

AUTHOR NOTE

We thank Geoff Patchin for the native English translation of CAIS. Correspondence concerning this article should be addressed to J. Willander, Stockholm University, Department of Psychology, Frescati Hagväg 14, SE-106 91 Stockholm, Sweden (e-mail: jwill@psychology.su.se).

REFERENCES

- ALLBUTT, J., LING, J., HEFFERNAN, T. M., & SHAFIULLAH, M. (2008). Self-report imagery questionnaire scores and subtypes of social-desirable responding: Auditory imagery, visual imagery, and thinking style. *Journal of Individual Differences*, *29*, 181-188.
- BENSAFI, M., PORTER, J., POULIOT, S., MAINLAND, J., JOHNSON, B., ZELANO, C., ET AL. (2003). Olfactomotor activity during imagery mimics that during perception. *Nature Neuroscience*, *6*, 1142-1144.
- BETTS, G. H. (1909). *The distribution and functions of mental imagery*. New York: Teachers College, Columbia University.
- BRETT, E. A., & STARKER, S. (1977). Auditory imagery and hallucination. *Journal of Nervous & Mental Disease*, *164*, 394-400.
- BULMER, M. G. (1979). *Principles of statistics*. New York: Dover.
- CARMINES, E. G., & ZELLER, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage.
- CLIFF, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin*, *103*, 276-279.
- DEVELLIS, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- EVANS, I. M., & KAMEMOTO, W. S. (1973). Reliability of the short form of Betts' questionnaire upon mental imagery: Replication. *Psychological Reports*, *33*, 281-282.

- FABRIGAR, L. R., WEGENER, D. T., MACCALLUM, R. C., & STRAHAN, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, **4**, 272-299.
- GALTON, F. (1880). Statistics of mental imagery. *Mind*, **5**, 301-318.
- GALTON, F. (1907). *Inquiries into human faculty and its development* (2nd ed.). London: Dent.
- GILBERT, A. N., CROUCH, M., & KEMP, S. E. (1998). Olfactory and visual mental imagery. *Journal of Mental Imagery*, **22**, 137-146.
- GISSURARSON, L. R. (1992). Reported auditory imagery and its relationship with visual imagery. *Journal of Mental Imagery*, **16**, 117-122.
- GORSUCH, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, **25**, 33-39.
- GORSUCH, R. L. (2003). Factor analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Volume 2. Research methods in psychology* (pp. 143-164). Hoboken, NJ: Wiley.
- HALPERN, A. R., ZATORRE, R. J., BOUFFARD, M., & JOHNSON, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, **42**, 1281-1292.
- HINKIN, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, **21**, 967-988.
- KOSSLYN, S. M., REISER, B. J., FARAH, M. J., & FLIEGEL, S. L. (1983). Generating visual images: Units and relations. *Journal of Experimental Psychology: General*, **112**, 278-303.
- MARKS, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, **64**, 17-24.
- MARKS, D. F. (1999). Consciousness, mental imagery and action. *British Journal of Psychology*, **90**, 567-585.
- MARKS, D. F., & ISAAC, A. R. (1995). Topographical distribution of EEG activity accompanying visual and motor imagery in vivid and non-vivid imagers. *British Journal of Psychology*, **86**, 271-282.
- McKELVIE, S. J. (1995). The VVIQ and beyond: Vividness and its measurement. *Journal of Mental Imagery*, **19**, 197-252.
- NIEMI, R. G., CARMINES, E. G., & McIVER, J. P. (1986). The impact of scale length on reliability and validity: A clarification of some misconceptions. *Quality & Quantity*, **20**, 371-376.
- O'CONNOR, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, **32**, 396-402.
- OKADA, H., MATSUOKA, K., & HATAKEYAMA, T. (2005). Individual differences in the range of sensory modalities experienced in dreams. *Dreaming*, **15**, 106-115.
- RECKASE, M. D. (1996). Test construction in the 1990s: Recent approaches every psychologist should know. *Psychological Assessment*, **8**, 354-359.
- RICHARDSON, A. (1977). The meaning and measurement of memory imagery. *British Journal of Psychology*, **68**, 29-43.
- SHEEHAN, P. W. (1967). A shortened form of Betts' questionnaire upon mental imagery. *Journal of Clinical Psychology*, **23**, 386-389.
- STREINER, D. L. (1998). Factors affecting reliability of interpretations of scree plots. *Psychological Reports*, **83**, 687-694.
- TABACHNICK, B. G., & FIDELL, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.
- TINSLEY, H. E. A., & TINSLEY, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, **34**, 414-424.
- VELICER, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, **41**, 321-327.
- VELICER, W. F., EATON, C. A., & FAVA, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41-71). Boston: Kluwer.
- WHITE, K. D., ASHTON, R., & BROWN, R. M. D. (1977). The measurement of imagery vividness: Normative data and their relationship to sex, age, and modality differences. *British Journal of Psychology*, **68**, 203-211.
- WORTHINGTON, R. L., & WHITTAKER, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *Counseling Psychologist*, **34**, 806-838.
- ZATORRE, R. J., HALPERN, A. R., PERRY, D. W., MEYER, E., & EVANS, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, **8**, 29-46.
- ZWICK, W. R., & VELICER, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, **99**, 432-442.

APPENDIX

The Clarity of Auditory Imagery Scale

English Version

Imagine the sounds listed below one at the time. Subjectively, how clearly do you hear the sounds of . . . (1 = not at all; 5 = very clear):

1. . . a clock ticking?
2. . . a car ignition?
3. . . a mobile phone ringing?
4. . . a dog barking?
5. . . water splashing from people bathing?
6. . . a person laughing?
7. . . the operation of a washing machine?
8. . . someone sneezing?
9. . . a group of people chatting?
10. . . dinner plates colliding?
11. . . the rustle of a bag?
12. . . a can of soft drink being opened?
13. . . birds singing?
14. . . the rustle of leaves under foot?
15. . . wind blowing in the trees?
16. . . paper being torn apart?

APPENDIX (Continued)

Swedish Version

Föreställ dig ett ljud i taget. Hur tydligt kan du för ditt inre föreställa dig ljudet av . . . (1 = *inte alls*; 5 = *mycket tydligt*):

1. . . . en klocka som tickar?
2. . . . bilar som startar?
3. . . . en mobiltelefon som ringer?
4. . . . en hund som skäller?
5. . . . plask från människor som badar?
6. . . . en person som skrattar?
7. . . . en tvättmaskin som tvättar?
8. . . . någon som nyser?
9. . . . sorlet från personer som pratar?
10. . . . tallrikar som slår emot varandra?
11. . . . prasslet från en påse?
12. . . . en burk läsk som öppnas?
13. . . . fåglar som kvittrar?
14. . . . löv som prasslar under fötterna?
15. . . . vinden som susar i träden?
16. . . . ett papper som rivs i tu?

(Manuscript received November 7, 2009;
revision accepted for publication January 30, 2010.)