Letter to the Editor

Factor Analysis in CAPD and the "Unimodal" Test Battery: Do We Have a Model that Will Satisfy?

Domitz and Schow (2000) present a series of exploratory factor analyses involving their Multiple Auditory Processing Assessment (MAPA) battery and the Screening Test for Auditory Processing Disorders (SCAN). Their interpretation of these analyses argues that four distinct aspects of central auditory processing are measured by these tests, but they note that some inconsistencies exist in their results. In a companion article, Schow, Seikel, Chermak, and Berent (2000) present results of a confirmatory factor analysis, which they suggest support the four-factor model. Herein, we point out several issues of contention with respect to the analysis and interpretation of these data. In addition, the lack of tests in other sensory modalities precludes the ability to demonstrate modality specificity with this battery, and therefore represents a major limitation with this proposed battery of tests. This and other issues are expanded on below.

Exploratory data analysis begins with basic descriptive statistics (measures of central tendency, dispersion, correlation, etc.) and often expands, as in this case, to more complex procedures and models. On the basis of descriptive statistics, relationships (correlations) between the subscales of the SCAN test (Keith, 1986; Schow and Chermak, 1999; Schow et al., 2000) and those obtained on the MAPA are rather low (Schow et al., 2000). One interpretation of this research (the view favored by Domitz and Schow, 2000) is that such tests measure distinct aspects of auditory processing ability. An alternative interpretation is that these tests are not reliable measures of a single underlying ability (Cacace and McFarland, 1995). Which interpretation do we accept?

There have been inconsistencies in reports of the factor structure of central auditory processing disorder (CAPD) test batteries. For example, Keith (1986) and Amos and Humes (1998) present results of factor analyses suggesting that measures from the SCAN test load on a single factor, whereas Schow and Chermak (1999) and Domitz and Schow (2000) present results suggesting that measures from the SCAN test load on two factors. However, several concerns exist for these two later reports that raise important issues and question the validity of their interpretation. These concerns include 1) the analytical methods used were not clearly stated, 2) an oblique rotation method was used that does not ensure independence among factors, and 3) the criterion used for accepting and retaining a factor (i.e., eigenvalues less than 1) is unacceptable by conventional standards. For example, in their work, Domitz and Schow (2000) simply state that they used factor analysis with an oblique rotation method. The description they provide is ambiguous because there are a number of different oblique rotation methods available that are based on different criteria (Harris, 1985). Moreover, the data included in these analyses were not the same in all studies. Schow and Chermak (1999) included scores from the Staggered Spondaic Word (SSW) test (Katz, 1968), whereas in those studies reported by Domitz and Schow (2000) and Amos and Humes (1998), SSW scores were not used.

The correlations between SCAN subscales used by Keith (1986), Schow and Chermak (1999), and Schow et al. (2000) are published, so we were able to reanalyze these data. In our reanalysis, data from 8- and 9-year-olds were used separately (Keith, 1986). In each case, we entered the published correlation matrix into the Statistical Analysis System (SAS, Hatcher, 1994) for examination. The FACTOR procedure, using principal components, was chosen as the extraction method, followed by retention of all factors with eigenvalues greater than 1. (These are the default options for the SAS FACTOR procedure and are perhaps the most common methods used.) In each of the four cases, only a single factor had a loading greater than 1 and therefore was retained in the analysis (Table 1). From Table 1, it can be seen that in all instances, each of the three scales had similar loadings on a single factor. These findings are also consistent with results reported by Amos and Humes (1998). Thus, consistent results emerge when the same analytical methods are applied to a correlation matrix constructed from the three SCAN subtests. Schow and Chermak (1999) obtained different results when principal components extraction and orthogonal varimax rotation were applied to a matrix including both SCAN and SSW test scores.

Domitz and Schow (2000) performed an analysis of only the SCAN subtests, but applied an *unidentified* oblique rotation method and accepted the second eigenvector, which had an eigenvalue of only 0.77. Conventionally, eigenvalues less than 1 are not retained for consideration (Harris, 1985). It is worth noting that when we retained two factors, regardless of the eigenvalues obtained, and used the Harris-Kaiser oblique rotation method (Hatcher, 1998), we obtained solutions similar to Domitz and Schow (2000) in three of the four cases (data not shown). Thus, some of the inconsistencies reported by these authors are caused by inconsistencies in the methods applied and do not appear to be caused by "chance characteristics of the data" as suggested by Schow et al. (2000).

Domitz and Schow (2000) also provided results of an analysis of eight subtest scores from their MAPA battery. They retained four factors, although they note that the

TABLE 1.

Study	FW	AFG	CW
Keith, 1986 (8-year-old data)	0.78	0.77	0.69
Keith, 1986 (9-year-old data)	0.77	0.80	0.64
Schow and Chermak, 1999	0.62	0.74	0.76
Domitz and Schow, 2000	0.73	0.77	0.67

FW, filtered words; AFG, auditory figure ground; CW, competing words.

TABLE 2.

Variable	Factor 1	Factor 2	Factor 3	Factor 4
SCAN-AFG	0.38	0.32	−0.57	−0.41
SCAN-FW	0.35	0.55	−0.34	0.02
SCAN-CW	0.74	0.03	0.16	−0.13
mSAAT-LE	0.60	0.53	0.07	0.03
mSAAT-RE	0.35	0.72	0.35	0.21
PP-LE	0.75	−0.30	−0.39	0.16
PP-RE	0.72	−0.28	−0.34	0.23
DD-LE	0.68	−0.23	0.25	−0.43
DD-RE	0.47	0.07	0.57	−0.19
CS-LE	0.51	−0.55	0.16	−0.17
CS-RE	0.48	−0.16	0.14	0.70
Eigenvalue	2.943	1.575	1.071	0.891

fourth factor had an eigenvalue of 0.896 and thus was not above the conventional noise level (Harris, 1985). However, it is worth noting that each of the four factors retained was composed mainly of just left- and right-ear scores from the same test. For example, the major loadings for the first factor were on left- and right-ear presentations of the pitch pattern (PP) test. Likewise, the major loadings for the second factor were on left- and right-ear presentations of the Selective Auditory Attention Test (SAAT; Cherry, 1980), and so on. One interpretation of these results is that left- and right-ear scores measure essentially the same underlying abilities and that these factors represent test-specific variance, rather than variance associated with some general auditory ability. It is also worth noting that the SAAT test was originally designed for binaural presentation. Domitz and Schow (2000) provide no evidence that ear differences on this test contributes any kind of meaningful information.

Given these methods, it is not surprising that a subsequent factor analysis of subtest scores for both the SCAN and MAPA yielded "some instability." In the companion article, Schow et al. (2000) use confirmatory factor analysis as an alternative means of evaluating their four-factor model. According to Becker (1990), "Authors should make every effort to identify equivalent models and to discuss whether such models offer plausible representations of the data." However, Schow et al. (2000) fail to consider alternative models. A fundamental problem with their approach is that several of the factors retained in their solution are defined in terms of only a single test (e.g., PPs presented to the left ear and PPs presented to the right ear). Anderson and Gerbing (1988) have suggested that multiple-indicator measurement models are preferable. This later approach allows one to separate variance associated with the trait in question from variance uniquely associated with a specific test. Ideally, the battery being analyzed would have three or four tests for each underlying trait.

We examined the data from Table 2 of Schow et al. (2000) by means of both exploratory and confirmatory factor analyses. We applied principal components analysis, only accepting eigenvalues greater than 1 (default values of the SAS FACTOR procedure). This analysis resulted in the three-factor solution shown in Table 2 (Factor 4 is also shown for completeness). The pattern of results is considerably different from that reported by Domitz and Schow (2000), who used an *unidentified* oblique rotation method. Again, this illustrates that the results are method dependent. It is also noteworthy that our first (largest) factor had positive loadings on all of the scales. We interpreted this as a general factor reflecting what is common to all of these tests. The other factors accounted for lesser amounts of variance and their interpretation is not at all clear. Factors 2 and 4 only have a single loading ≥0.7 and Factor 3 has none.

We next used the CALIS procedure from SAS (Hatcher, 1994), with the maximum likelihood method, to replicate both the model evaluated by Schow et al. (2000) and one alternative model. For both models, we estimated the variance associated with the underlying factors and test-specific error. To replicate the Schow et al. (2000) model, we included filtered words (FW), auditory figure ground (AFG), monaural Selective Auditory Attention Test, left ear (SAAT-LE), and monaural Selective Auditory Attention Test, right ear (mSAAT-RE) on the first factor; dichotic digits, left ear (DD-LE) and dichotic digits, right ear (DD-RE) on the second factor; pitch patterns, left ear (PP-LE) and pitch patterns, right ear (PP-RE) on the third factor; competing sentences, left ear (CS-LE), competing sentences, right ear (CS-RE), and competing words (CW) on the fourth factor. Additionally, we estimated the covariance between these factors.

Our replication of the Schow et al. (2000) model provided a goodness-of-fit index (GFI) of 0.85 (compared with their reported GFI of 0.86) and an adjusted goodness-of-fit index (AGFI) of 0.72 (compared with their reported AGFI of 0.77). These values are reasonably close to the values presented by Schow et al. (2000) considering that these authors did not identify the specific procedure they used. The alternative model we used considered a single common factor on which all tests were included and four specific factors on which each pair of left- and right-ear presentations were included (i.e., mSAAT-LE and mSAAT-RE, PP-LE and PP-RE, DD-LE and DD-RE, CS-LE and CS-RE). Our orthogonal model, with a general factor and four test-unique factors, generated a GFI of 0.84 and an AGFI of 0.72. These values are also reasonably close to the values obtained by the Schow et al. (2000) model. Therefore, we conclude that the data are consistent with both models, although neither provides a good fit of the data. Thus, these tests may measure four distinct abilities, or a single ability and several test-unique effects.

A major limitation with all of these studies and analyses considered herein centers on their potential inability to identify auditory-specific effects. That is, testing protocols advocated by Domitz and Schow (2000) included tests that only involved *auditory* stimuli. Hence, it is not possible to determine the extent to which these tests assess modality-specific (perceptual) abilities as opposed to more general supramodal (cognitive) abilities (Cacace and McFarland, 1998; McFarland and Cacace, 1995). Inclusion of comparable tests in different sensory modalities is necessary to determine whether MAPA and SCAN measure perceptual abilities, as they are purported to do. That is, if they are valid measures of auditory perception, they should not predict performance on tests in other sensory modalities, such as tactile or visual tests. The importance of having a battery of tests in multiple sensory modalities that can delineate auditory-specific effects has been emphasized in a recent consensus statement on the topic of auditory processing disorders in children (Jerger and Musiek, 2000). As it stands, we cannot rule out the possibility that the tests in question are actually measures of cognitive abilities. In addition, it remains to be seen whether these tests correlate with other indices of auditory processing. This brings up a second important issue that was not addressed by these authors. That is, what do these tests correlate with and what do they predict? In school-age children, this could include factors that may be associated with intelligence, spelling ability, reading ability, math skills, attentional skills, distractibility, musical ability, and so forth.

Given these findings, we conclude that the MAPA battery is not a reliable index of a single auditory ability because individual tests scores on the MAPA battery do not correlate highly (with the exception of right- and left-ear presentations of the same scales). Schow et al. (2000) suggest that these scales measure a number of distinct abilities. However, to support this claim, it will be necessary to demonstrate that these tests correlate with other indices of these separate functions. That is, Schow et al. (2000) now have the task of validating four measures. This involves showing that the tests in question correlate with what they are theoretically expected to correlate with (e.g., other measures of the purported auditory ability). In addition, this validation process involves demonstrating modality specificity by showing that these tests do not correlate with tests in other sensory modalities.

Dennis J. McFarland
The Wadsworth Center, Albany, NY

Anthony T. Cacace
Albany Medical College, Albany, NY

References

Amos, N. E., & Humes, L. E. (1998). SCAN test-retest reliability for first- and third-grade children. *Journal of Speech, Language and Hearing Research, 41,* 834–845.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411–423.

Becker, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, 107, 260–273.

Cacace, A. T., & McFarland, D. J. (1995). Opening Pandora's box: The reliability of CAPD tests. *American Journal of Audiology, 4*, 61–62.

Cacace, A. T., & McFarland, D. J. (1998). Central auditory processing disorder in school-aged children: A critical review. *Journal of Speech, Language and Hearing Research, 41,* 355–373.

Cherry, R. S. (1980). *Selective Auditory Attention Test*, St. Louis: Auditec.

Domitz, D. M., & Schow, R. L. (2000). A new CAPD battery-Multiple Auditory Processing Assessment: Factor analysis and comparison with SCAN. *American Journal of Audiology, 9,* 101–111.

Harris, R. J. (1985*) A Primer of Multivariate Statistics,* 2nd ed., Academic Press, Orlando.

Hatcher, L. (1994). *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute.

Jerger, J., & Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorders in school-aged children. *Journal of the American Academy of Audiology, 11,* 467–474.

Katz, J. (1968). The SSW test: An interim report. *Journal of Speech and Hearing Disorders, 33,* 132–145.

Keith, R. W. (1986). *SCAN: A Screening Test for Auditory Processing Disorders*. San Antonio, TX: The Psychological Corporation, Harcourt, Brace, Jovanovich, Inc.

McFarland, D. J., & Cacace, A. T. (1995). Modality specificity as a criterion for diagnosing central auditory processing disorders. *American Journal of Audiology, 4,* 36–48.

Schow, R. L., & Chermak, G. (1999). Implications from factor analysis for central auditory processing disorders. *American Journal of Audiology, 8,* 137–142.

Schow, R. L., Seikel, J. A., Chermak, G. D., & Berent, M. (2000). Central auditory processing and test measures: Asha 1996 revisited. *American Journal of Audiology, 9,* 63–68.