Critical Review:

Are currently available pre-packaged behavioural test batteries (SCAN and MAPA) effective for use in the assessment and or diagnosis of Auditory Processing Disorder (APD) in children assuming the American Speech-Language Hearing Association (ASHA) definition of APD?

Bonnie Lampe
M.Cl.Sc (AUD) Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This literature review investigates the effectiveness of currently available pre-packaged behavioural test batteries in the assessment and diagnosis of APD in children. This review examines the effectiveness of both the SCAN and MAPA pre-packaged behavioural test batteries for children in relation to the ASHA (2005) definition of APD and APD assessment. These assessment batteries are assessed in terms of test-retest reliability, test condition effects, normative data, factor analysis and efficiency in terms of the number of tests in each battery and the cost effectiveness. Study designs include: within groups repeated measures, non-randomized cohort studies, non-randomized case-control study, single group test, and single group factor analysis. Overall, current research suggests that the SCAN alone should not be used for assessment and diagnosis of APD in children and that the MAPA test battery shows promise. However, with limited peer-reviewed and published literature available on the MAPA, definitive conclusions cannot be made regarding the utility of the MAPA for assessment and diagnosis of APD in this population.

Introduction

The diagnosis and treatment of Auditory Processing Disorders (APD) is at the forefront of research in audiology. Currently there are two pre-packaged test batteries available on the market to assess and diagnose APD in children. These batteries are: the Screening Test for Auditory Processing Disorder (SCAN) and the Minimal Auditory Processing Assessment (MAPA). The purpose of these test batteries is to allow for early diagnosis and detection of APD in children in order to augment early intervention. However, if these test batteries do not effectively achieve their purpose, a misdiagnosis of APD is possible. A misdiagnosis of APD, especially for a child, can often be detrimental socially and emotionally and can result in a drain on already scarce educational resources.

ASHA (2005) defines APD in terms of a deficit in the neural processing of auditory stimuli. In particular, ASHA (2005) suggests that this deficit is demonstrated by poor performance in one or more of six areas: (1) auditory pattern recognition; (2) temporal processing; (3) auditory performance with degraded speech signals; (4) auditory performance with competing acoustic signals; (5) auditory discrimination; (6) localization/lateralization. In children, APD can be associated with difficulties in learning, speech and language. Also, children with APD often have difficulties behaviourally, emotionally and socially. These manifestations, however, are not exclusive to APD. This fact makes the assessment of APD extremely challenging.

In terms of assessment, ASHA (2005) recommends that a patient-specific, multi-disciplinary team-based approach be used. As such, clinicians need to be aware of all the assessment procedures available and what they assess in order to ensure the correct test battery is chosen for the individual patient. ASHA (2005) offers a guide to clinicians in terms of the available testing types that can be included in a test battery. These test types are: auditory discrimination tests; auditory temporal processing and patterning tests; dichotic speech tests; monaural low-redundancy speech tests; binaural interaction tests; electroacoustic measures and electrophysiological measures. Research by Katz et al. (2002) suggests that behavioural testing is most appropriate for children as there is little evidence to support APD objective testing methods for use with children and there is limited availability and high expense associated with the tools required for these measures.

The SCAN is the most widely used pre-packaged test battery in APD assessment. The SCAN is composed of three subtests: filtered words (FW), auditory figure-ground (AFG) and competing words (CW). All test types are administered using monosyllabic real words. SCAN administration takes about 20 minutes. More recently the SCAN-3 has been released which includes an additional test for temporal resolution. The SCAN can be scored by adding up the total correct responses in each section (subtest) of the test as well as looking at the overall (composite) score. The subtest scores and composite scores are then compared with normative data and interpretation of results in made, including percentile ranking tables (Keith, 1986).

The MAPA is a new battery and is composed of three domains of behavioural testing that are deemed by Chermak (2001) to be the most relevant for APD screening: auditory pattern temporal ordering (APTO), binaural integration and binaural separation (BIBS) and monaural separation closure (MSC). The MAPA includes tests of: Monaural Selective Auditory Attention (mSAAT); Tap test; Pitch Patterns (PP); Dichotic Digits (DD) and Competing Sentences (CS) divided into 8 subtests. The MAPA test materials involve test segments that are either tonal in nature or are composed of recorded speech (Musiek & Chermak, 2007). The MAPA can be administered in 21 minutes. Scoring is similar to that of the SCAN with a simple tallying of test scores. Incomplete normative mean data is available for the latest MAPA version 1.0 (2007). MAPA 1.0 (2007) has recently been released to the public and is available for purchase from AUDiTEC of St.Louis.

With the introduction of the new MAPA battery for APD assessment, there is a need for a comparison between the widely utilized SCAN and the MAPA. Additionally, it becomes important to evaluate the effectiveness of both such test battery approaches in the assessment and diagnosis of APD in children. With the lack of a gold-standard definition of APD in children, the ASHA (2005) definition will be used for this evaluation.

Objectives

The primary objective of this literature review is to critically evaluate the effectiveness of the SCAN and MAPA for the assessment of APD in children. In particular, this reveiw will relate these two test batteries to the ASHA (2005) definition of APD by critically assessing the effectiveness of the test battery measures in the diagnosis of APD in children.

Methods

Search Strategy

Computerized databases including Medline and PubMed were searched using the following strategy: auditory processing disorder (OR) central auditory processing disorder (OR) CAPD (OR) APD (AND) SCAN (AND) MAPA (AND) test battery. In addition to using these computerized search engines and keywords, the reference lists were reviewed and pertinent articles were sought out. This search was limited to articles that were written in English between 1997 and 2010.

Selection Criteria

Studies included in this critical review were required to examine the effectiveness of the SCAN and/or MAPA when administered to the school-aged population. In all of the studies selected, children were considered to be at risk for APD based on academic performance, namely language and reading difficulties. No limits were set on the demographics of the children and studies from North America and the United Kingdom were included.

Data Collection

Results of the literature search yielded ten articles consistent with the selection criteria: three within groups repeated measures studies, three non-randomized cohort studies, one single group test, and three single group factor analyses. These studies yielded levels of evidence of III (Amos & Humes, 1998; Summers, 2003; Conlin, 2003; Emerson et al., 1997; Schow & Chermak, 1999; Domitz & Schow, 2000; Schow et al.,2000) and II (Marriage et al., 2001; Dawes & Bishop, 2007; Singer et al., 1998) based on Dollaghan (2007). Although there is a diverse group of studies and researchers presented in this literature review, there are a few researchers that are common amoung several of the studies: three studies with Gail D. Chermak (Washington State University) and three studies with Ronald L. Schow (Idaho State University).

Results

Test-Retest Reliability

Amos and Hurnes (1998) used a within groups repeated measures experimental design to assess the test-retest reliability of the SCAN when administered to a group of children twice: initially for the first time and then again six to seven weeks later. Participants included a total of 47 children aged 6 to 9 years old (25 first graders, 22 third graders). Participants were all Caucasian with English as their primary language and were considered to be performing at an age appropriate academic level by their teachers.

Subtest scores, composite scores, percentile rank and age-equivalent outcomes of this study indicated that all subjects performed better on the second administration of the SCAN with the exception of the AFG subtest. Unfortunately, however, no quantifiable relationship was found between the first and second test administration when Pearson r test-retest correlations and scatter plots were generated. Further data analysis of the variance (ANOVA) between the scores of the first and third graders was conducted to assess the potential effects of age or grade on the subtest scores, composite scores, percentile rank and age-equivalent scores. Results demonstrated that third graders had significantly higher FW raw scores than first graders suggesting a possible maturation effect (Amos & Humes,1998).

In an attempt to alleviate any test-retest reliability issues for the MAPA, designers included two equivalent tests under each test category resulting in two separate MAPA test forms, A and B. In 2003, Conlin used a within groups repeated measures experimental design to ensure that forms A and B were indeed equivalent. The two forms of the MAPA were administered to a total of 48 children. Forms A and B were found to correlate well under each category. Additionally in 2003, Summers further evaluated the test-retest reliability of the two form MAPA using a within groups repeated measures experimental design. The MAPA was administered to 19 children aged 8 – 11. Results demonstrated that the test-retest reliability of the MAPA meets the generally accepted standard of reliability ($r \geq 0.7$) (Summers, 2003). Therefore, this two form design utilized by the MAPA seems to help alleviate any issues related to test-retest reliability.

Testing Condition Effects

Emerson, Kami, Seikel and Chermak (1997) conducted a single group test with 6 children (5 girls, 1 boy), looking at the effect of ambient noise during APD testing. All subjects were administered the SCAN in both school and booth settings with one week between the initial and second presentation of the SCAN to each child. The presentation of the SCAN in the booth or in the school setting was alternated to avoid order effects. Results revealed depressed scores (5 out of 6) for the SCAN administered in the school setting compared to the booth especially for the AFG subtest. Based on the diagnosis criteria for the SCAN, none of the subjects would be identified as being at risk for APD by the SCAN when administered in the booth. When the SCAN was administered in the school setting however, 2 out of 6 children scored below criterion on the AFG subtest and overall composite score. The mean performance differences between sites were 7.2 for composite score and 3 for the AFG subtest. The authors suggest that, for this reason, further normative values should be determined for the SCAN administered in a school/noisy setting.

Normative Data Relevance

Marriage, King, Briggs and Lutman (2001) conducted a non-randomized cohort study to compare and identify whether the published normative values for the SCAN are accurate and useable for children outside the United States and Canada. 133 children from the UK were included in this study. Results suggest that the published normative values for the SCAN test are not valid for direct application for children from the UK (Marriage, 2001).

To further these results, Dawes and Bishop (2007) conducted a non-randomized cohort study comparing SCAN results from a group of 99 children from the UK (more representative of the UK population in terms of demographics than the 2001 study) to published SCAN normative data. Results were found to be the same as the 2001 study, suggesting that the SCAN should be administered with caution for populations outside the US and Canada (Dawes & Bishop, 2007).

Test Battery : effectiveness and efficiency

Singer, Hurley and Preece (1998) conducted a non-randomized cohort study with a between groups design to determine the efficacy of a test battery versus a single test approach to APD assessment. A total of 238 children ranging in age from 7 to 13 years were divided into two groups: normal learning (NL, n = 91) and classroom learning disabled (CLD, n= 147). Bilingual children and children suspected of attention deficit were excluded from this study. A total of 7 APD tests, found to be independent of one-another, were administered to all subjects. To assess the ability of these tests to separate the children with CLD from children in the NL group, data was analyzed using clinical decision analysis (CDA) procedures including sensitivity measures, specificity measures, A' and receiver operating curves (ROC). Additionally, posterior probabilities were calculated for the following conditions: membership in the CLD group given a positive test result and membership in the NL group given a negative test result. Lastly, the cost effectiveness of a test battery approach when compared to a single test approach was assessed (Singer et al., 1998).

The largest A' value for a single test administration was found to be 0.90 (BFT alone) with a total cost of identification of $112.56 per person, while the largest A' value calculated overall was 0.93 (BFT*FST*MLD OR BFT±FST OR BFT±MLD) at a total cost of identification being $317.69 per person. The highest positive posterior probabilities were calculated for a two test (BFT+ MLD) approach at 87%. Although this two test approaches yielded an A' value of 0.89 the hit rate was found to be only 59%. Cost effectiveness analysis demonstrated that a two test protocol had twice the per identification cost with little improvement in hit rate (Singer et al., 1998).

The authors concluded that there is no single test or test battery that optimizes each criterion and that multiple tests do not necessarily provide increased identification power when compared to a one or two test sequence. As such, with the increased cost associated with multiple tests, these authors recommend the use of an abbreviated battery (with two tests) in the identification of APD (Singer et al., 1998).

Factor Analysis

Schow and Chermak (1999) conducted a retrospective single group exploratory factor analysis of the SCAN with the purpose of identifying the nature of auditory processes probed by this test battery. The data analyzed in this study was from a group of 331 school-aged children aged 6 to 17 referred for APD testing due to underachievement, poor classroom performance and/or attention limitations.

Factor analysis for the SCAN was conducted using rotated factor loadings of the rotated solution for each variable. The orthogonal solution with principal components factoring was used as it provided the simplest solution and assumed complete independence. Two factors emerged from this analysis: a binaural separation/competition factor (loaded on CW) and a composite monoaural low-redundancy degradation factor (loaded on AFG and FW). The authors concluded that the three subtests of the SCAN only look at two out of the six factors associated with the APD definition proposed by ASHA (2005).

Domitz and Schow (2000) conducted a single group exploratory factor analysis of both the SCAN and MAPA and performed exploratory factor analysis for each. Final data was based on 81 children (40 boys and 41 girls) aged 8.8 to 9.9. For the SCAN, findings were similar to that of Schow and Chermak (1999). For the MAPA, four factors emerged: auditory pattern/temporal ordering (APTO), monaural separation/closure (MSC), binaural separation (BS) and binaural integration (BI).

The above data from Domitz and Schow (2000) was reanalyzed by Schow et al. (2000) using confirmatory factor analysis (CFA) to add more significance to the results. For this CFA, a model was specified in which each test loaded on one of the four factors determined by Domitz & Schow (2000). A comparison was made between the subtest correlation matrix and the actual correlation matrix using a χ^2 test. The χ^2 test indicated a significant difference between the two matrices (χ^2 = 81.36, $p < 0.001$). However, due to the conservative nature of the χ^2 test and the large number of variables in this study, a ratio of χ^2 to degrees of freedom was looked at as well to see if this resulted in a better fit. Results from this ratio demonstrated that this model can in fact be categorized as good. The authors conclude that the four-factor model proposed here generates fit indices approaching good fitting levels. The authors suggest that the eleven subtests looked at in this study represent all four factors suggested by Chermak (2001) to be the most important for behavioural assessment of APD in children.

Additionally, for the MAPA, Summers (2003) completed additional single group factor analyses (both exploratory and confirmatory) on test results from a separate group 119 children. Results further demonstrate that the underlying factors are indeed loaded on the auditory domain, specifically APTO, BIBS and MSC.

Discussion

According to this literature review, the SCAN has many underlying weaknesses. For one, test-retest reliability is fairly poor. Second, the SCAN is sensitive to the administration environment. Third, scores are highly dependent on the comfort level of the child with American English as opposed to British English. Fourth, factor analysis suggests that the SCAN accounts for only two out of six factors outlined by the ASHA (2005) definition, suggesting limitations of the SCAN in terms of proper assessment and diagnosis of children at risk for APD. With this said however, any additions to the SCAN should be made with caution as Singer et al. (1998) concluded that, with the addition of tests into a test battery, overall cost of test battery administration can outweigh the improved effectiveness. Alternatively, from the limited resources available, the MAPA shows much promise in terms of fulfilling the guidelines set-out for the behavioural assessment and diagnosis of APD in children. When considering all of the results presented in this paper, however, one should acknowledge some procedural short-comings. These shortcomings are presented below.

In the Amos and Hurnes (1998) study looking at the test-retest reliability of the SCAN, the authors neglected to take into account the client-clinician rapport factor. An alternative reason for the improved scores on the second administration of the SCAN could be attributed to the established rapport between the child and clinician as the clinician remained the same for both administrations.

In the study presented by Emerson, Kami, Seikel and Chermak (1997) looking at the effect of background noise, there are two major criticisms. First, the participant pool is quite small and is composed of significantly more female participants than male participants. As such, the reliability of this study is very low. Second due to the concern over the test-retest reliability associated with the SCAN, inaccurate results could have been achieved.

In the Domitz and Schow (2000) and Schow and Chermak (1999) studies, exploratory factor analyses of both the SCAN and MAPA were conducted with several limitations. For one, in the Domitz and Schow (2000) study, subjects were from a very small age group range (8.8 to 9.9). An assessment battery must have proven effectiveness over a much larger age range and as a result, more analysis should be done with participants that span a larger age range. Additionally, in terms of all presented factor analysis studies, exploratory factor analysis (EFA) was used. EFA looks at factors from an infinite number of possible solutions and there is an increased probability that the final solution will be due to chance. Additionally, an exploratory model does not delineate how well the data actually fits into the solution, it simply derives a solution. Instead an additional confirmatory factor analysis was undertaken in the same year by Schow et al. (2000). Results suggest good correlation but when these authors were assessing the robustness of the model, they used the χ^2 test and ratio which may not have been the best statistical methods to use. Results from the χ^2 test were limited because the model was large and attempted to reproduce a matrix with 55 indispensible correlations. Instead it seems as though the authors struggle to conclude that their model is significant by looking at the χ^2 ratio and in the end come to a somewhat hesitant conclusion that these results would suggest that the model is significant. Perhaps instead, the authors could have used a G- test to analyze this data.

For the Singer et al. (1998) study, which looks at the cost-effectiveness of a test battery approach versus a single test, research is flawed in that cost and factor structure are the only issues attributed to efficiency and effectiveness. In this field, there are many more factors that come into play when evaluating effectiveness and efficiency of an assessment procedure (i.e. how enjoyable the task is to the patient) and these results should be taken lightly.

With regard to the two theses presented in this paper, neither Summers (2003) and Conlin (2003) are peer reviewed which significantly limits their credibility.

A consistent and very significant limitation that can be applied to every study presented in this literature review stems from the lack of a gold-standard definition of APD. Participants in each study were grouped in the possible APD category based mainly on sub-par academic performance. Although poor academic performance is thought to be consistent with APD in children, there are many other factors that can also contribute to poor academic performance. Additionally, children placed in the control, or "normal" groups in the various studies presented here could have in fact had auditory processing difficulties that were not yet manifested in the classroom.

Conclusion and Clinical Implications

There are clearly many challenges when it comes to the assessment of APD in children. From the studies summarized in this literature review it becomes clear that the SCAN is limited in terms of the ASHA definition for APD diagnosis. The MAPA seems promising but has only recently been released to the public and is deemed to still be a work in progress. This literature supports the ASHA (2005) definition which suggests that the assessment for APD should be more individualized as opposed to generic, as a test battery approach would allow. It is reasonable to conclude that currently available pre-packaged test batteries for the assessment of APD are not as effective as more individualized case-by-case assessment protocols.

Future research should include a search for objective evidence to support a gold standard definition of APD. Additionally, perhaps research should concentrate on developing a single task, using all possible examination modalities, to assess APD in children in order to optimize the clinician's time.

References

American Speech-Language-Hearing Association. (1996). Central auditory processing: current status of research and implications for clinical practice. *American Journal of Audiology*, 5, 41 – 54.

American Speech-Language-Hearing Association. (2005). (Central) auditory processing disorders – The role of the audiologist. Position Statement of the Working Group on Auditory Processing Disorders of the American Speech-Language-Hearing Association. Rockville, MD: Author.

Amos, N.E. & Hurnes, L.E. (1998). SCAN test-retest variability for first- and third-grade children. *Journal of Speech, Language & Hearing Research,* 41(4), 834 – 846.

Chermak, G. D. (2001). Auditory processing disorder: An overview for the clinician.*The Hearing Journal*, 54(7), 10-21.

Conlin, L. (2003). Form equivalency on the Beta III version of Multiple Auditory Processing Assessment (MAPA). Master's thesis, Idaho State University, Pocatello, ID.

Dawes, P. & Bishop, D.V.M. (2007). The SCAN-C in testing for auditory processing disorder in a sample of British children. *International Journal of Audiology*, 46, 780-786.

Dollaghan, C. (2007). The handbook of evidence-based practice in communication disorders. Baltimore: Paul H. Brookes Publishing Co.

Domitz, D.M. & Schow, R.L.(2000). A New CAPD Battery-Multiple Auditory Processing Assessment: Factor Analysis and Comparisons with SCAN. *American Journal of Audiology*, 9(2), 101 – 112.

Emerson, M.F., Kami, K.C., Seikel, J.A. & Chermak, G.D. (1997). Observations on the use of SCAN to identify children at risk for Central Auditory Processing Disorder. *Language, Speech and Hearing Services in Schools*, 28, 43 -49.

Katz, J., Johnson, C.D., Brandner, S., Delagraange, T., Ferre, J, King, J. Kossover-Wechter, D., Lucker, J., Medwetsky, L., Saul, R., Rosenberg, G.G., Stecker, N.,& Tillery, K. (2002). Clinical and research concerns regarding the 2000 APD consensus report and recommendations. Audiology Today.

Keith, R. (1986). SCAN- A Screening Test for Auditory Processing Disorders. San Diego, CA: The Psychological Corporation.

Marriage, J., King, J., Briggs, J. & Lutman, M.E. (2001). The reliability of the SCAN test: results from a primary school population in the UK. *British Journal of Audiology*, 35(3), 199-208.

Musiek, F. E. & Chermak, G.D. (2007) [Chapter 6] Handbook of (central) auditory processing disorder: From science to practice (Vol. 1),(Central) auditory processing disorder:Auditory Neuroscience and Diagnosis. San Diego: Plural.

Schow, R.L. & Chermak, G.D. (1999). Implications from Factor Analysis for Central Auditory Processing Disorders. *American Journal of Audiology*, 8 (2), 137-142

Schow, R.L., Chermak, G.D., Seikel, J.A., Brockett, J.E., & Whitaker, M.M. (2007). Multiple Auditory Processing Assessment. St.Louis, MO: Auditec.

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

Singer, J., Hurley, R. M. and Preece, J. P. (1998). Effectiveness of Central Auditory Processing tests with children. *American Journal of Audiology,* 7(2), 73 – 84.

Summers, S. A. (2003). Factor structure, correlations, and mean data on Form A of the Beta III version of Multiple AuditoryProcessing Assessment (MAPA). Master's thesis, Idaho State University, Pocatello, ID