Critical Review:
What is the validity of using nonverbal intelligence scores as an exclusionary criterion when identifying children with specific language impairment?

Katherine Harder
M.Cl.Sc (SLP) Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This paper is a critical review of the existing literature examining the validity of nonverbal intelligence as an exclusionary criterion used to categorize children as having specific language impairment. Following the completion of a computerized database search seven articles meeting the inclusion criteria were included for analysis. Study designs include: nonrandomized control trials (between groups and mixed), within group (repeated measures), and expert opinion articles. Overall, research does not support the use of nonverbal intelligence scores as an exclusionary criterion used to categorize children as having specific language impairment. A discussion of clinical recommendations is included.

Introduction

Specific Language Impairment (SLI) is a significant impairment in one's ability to use language (either expressive or receptive) that occurs in the absence of hearing impairment, mental retardation, motor disorder, socioemotional dysfunction, and/or neurological damage (Friel-Patti, 1999; Plante, 1998).

Children with SLI represent a unique clinical group that has been difficult to identify. One of the first criteria aimed at defining this group was to rule out those whose language difficulties were due to general cognitive impairment (Plante, 1998). This criterion was operationalized as a score on a nonverbal test of intelligence within one standard deviation of the standardized mean (>85; Stark & Tallal, 1981).

In recent years, variations and adaptations of this method have been used by clinicians and researchers in the field of speech-language pathology to identify children with SLI, as a result much heterogeneity exists in the way that nonverbal IQ scores are used to identify children with SLI; such that there currently is not one universally accepted method for identifying children with SLI (Plante, 1998; Krassowski & Plante, 1997). Current research attention has focused on the way that nonverbal intelligence is used to identify children with SLI (Rice, 2004).

Studies have demonstrated that SLI is an enduring disorder effecting children during the school years and continuing throughout the lifespan (Friel-Patti, 1999). This then has far reaching implications on a person's quality of life including educational and occupational success. Ensuring that accurate criteria are established and used when labeling children with SLI is important since eligibility for services is determined based on diagnosis. Therefore, determining the appropriateness of using nonverbal intelligence scores as a criterion for determining the presence or absence of a language impairment is of critical importance for ensuring that children with language impairment are provided with the treatment they need.

Objectives

The primary objective of this paper is to review and critically evaluate the existing literature concerning the use of nonverbal intelligence scores as a diagnostic criterion for SLI in children. A secondary objective is to discuss the clinical implications of these findings as they relate to the field of speech-language pathology and to propose recommendations for speech-language pathologists who are working with children with SLI. This information is also relevant to researchers who use nonverbal intelligence scores as a criterion for grouping participants into experimental groups.

Methods

Search Strategy

Computerized databases were searched including CINAHL and PsychInfo using the following search parameters: (cognitive referencing) AND (specific language impairment) OR (language disorder); (specific language impairment) OR (language disorder) AND (nonverbal intelligence) OR (nonverbal IQ) OR (performance IQ). Articles were also located by searching the reference section of relevant articles.

Selection Criteria

Articles were selected to be included in this paper if a main focus of the study was on the use of nonverbal intelligence scores in order to classify children as specific language impaired. Commentaries on

identifying children with specific language impairment were also included. Only articles written in English were selected and no date restriction was used.

Data Collection

Upon completion of the previously described literature search the following articles met criteria to be included in this analysis: three nonrandomized clinical trials (two between groups and one mixed) as well as two within groups (repeated measures). In addition, two expert opinion articles were included. These articles provide an overview of the current state of research in the area of selection methods for children with specific language impairment.

Results

The following studies are presented in groups according to the type of study design utilized.

Nonrandomized Clinical Trial: Between Groups

The purpose of the study by Stark and Tallal (1981) was to establish a standard method for the selection of children with SLI. One of the criterions used in this study was a nonverbal IQ cutoff score of 85. To determine each participant's nonverbal IQ score, either the WISC-R or WPPSI, both widely accepted intelligence tests was used. As a secondary procedure, a subset of the participants were administered a second nonverbal intelligence test with nonverbal instructions (Hiskey-Nebraska Test of Learning Aptitude). Low performance across both measures confirmed poor nonverbal IQ, which was interpreted as ruling out language influences on nonverbal intelligence. Using a between groups design they recruited participants between the ages of 4 to 8.5 years. Children with normal hearing, intelligence, and social and emotional adjustment were referred. The researchers completed a test battery with each child to see if they met criteria for SLI. The results of the testing revealed that of the 132 children originally referred by clinicians 39 met full criteria for SLI. This indicates that some of the children who were considered by clinicians to have language impairment did not meet criteria for SLI. Of those children some did not meet criteria based on their score on the nonverbal intelligence measure. Of the 90 normal language children, 47 fully met criteria for not being SLI. The groups were matched for age, performance IQ, and socioeconomic status but not participant sex (control group: male = female; experimental group: 2.45:1).

This study has level 2b evidence. The study outlines how the experimental group was recruited; however, one weakness of the study is that it does not state how the control group was recruited. The tests used to identify and compare children in each group were clearly outlined and a rationale for each test used was provided. One weakness was that the ten children with the lowest nonverbal intelligence scores, rather than a randomly chosen sample of children, were administered an additional nonverbal intelligence test that used nonverbal instructions. The statistical analyses completed in this study include: descriptive statistics (e.g. mean and standard deviation) and correlation analyses between tests. These analyses were appropriate for the purpose of the study. Overall, the study provides suggestive evidence that clinically identified groups have considerable variability with nonverbal intelligence.

The primary goal of Webster et al (2006) was to evaluate language, nonverbal cognitive, and motor development in children with developmental language impairment (henceforth, SLI). The experimental group consisted of school age children with SLI (n=11) and a control group (n=12) consisted of children with normal language but headache disorder. All children completed tests of language, nonverbal cognition, and motor development. Data analysis included t-tests, Mann Whitney (where appropriate), and chi-square tests. Results revealed lower nonverbal intelligence scores in the SLI group, although both groups scored in the normal range.

This study has a 2b level of evidence. A weakness of the study is a small sample size (n=23). In addition, the control participants were not randomly selected which may introduce a bias. On the other hand, this could be a positive since both the control and experimental groups have similar experiences (i.e., hospital visits). Another weakness is that the control group may have been "super controls" since their scores were significantly higher than the norms for the test. A strength of the study is that the testers were qualified therapists who were blinded to the participants groupings. The data analyses were appropriate. Overall, the results provide borderline evidence that children with SLI score lower than groups without language impairment on tests of nonverbal IQ.

Nonrandomized Clinical Trial: Mixed Group

The purpose of the study by Miller and Gilbert (2008) was to compare performance of the same children with and without language impairment on two nonverbal intelligence tests. They selected 204 grade eight students to participate. Children were classified into one of four groups based on their language and nonverbal IQ scores. The two groups of relevance to the present review are the typically developing group (scored above language and nonverbal IQ cutoffs) and the SLI group (scored above nonverbal IQ, but below language

cutoff). Each child was administered two nonverbal intelligence tests. A cutoff score of 85 on the nonverbal intelligence test was used.

Results of the mixed model ANOVA indicated a significant main effect of group, a significant main effect of test, and a group by test interaction. No significant difference between the mean test scores was found for the typically developing group. Scores on the two tests were significantly correlated for the typically developing group. In contrast, a significant mean difference was found between the two tests for the SLI group. A post hoc analysis (unequal N HSD) showed no difference between the two tests for the typically developing group, but that the SLI group scored significantly higher on the WISC-III than on the UNIT. Based on the scores obtained on the UNIT, 16% of SLI children were reclassified as typically developing and 74% of typically developing children were reclassified as SLI.

This study is a 2b level of evidence. The method of selection and grouping of participants was well described in the article. A strength of the study is that standard administration procedures were followed by trained examiners for each test. The statistical analyses were appropriate and well described. Overall, the results provide compelling evidence that the same child may be classified as SLI on one test of nonverbal IQ and as typically developing on another test of nonverbal IQ.

Within Group (Repeated Measures)

A study by Dethorne and Watkins (2006) examined the extent that language abilities are associated with nonverbal IQ in children with language impairment. Children between the ages of 4 to 8 years with language impairment who were currently being seen by a speech-language pathologist participated in this study (n=30). All participants met common criteria, which were well defined. The authors chose not to use an IQ cutoff score as an inclusion criterion. Each child was administered the same battery of tests consisting of a standardized IQ measure, standardized language tests, and criterion-referenced language measures. Data analysis included descriptive statistics and partial correlations. Results of the study indicate that nonverbal IQ had moderate-high associations with semantic and morphosyntactic abilities when assessed using standardized measures. No significant associations were found between nonverbal IQ and language when criterion-referenced language measures were used. Analysis of individual difference scores between language and nonverbal IQ found that for some children language exceeded nonverbal IQ and in others nonverbal IQ exceeded language.

This study is a 2c evidence level. Participants consisted of 8 females and 22 males, this division seems reasonable given that more males than females have language disorders in the general population. One weakness is that the study did not have a control group with which to compare the results of the experimental group. The authors provided appropriate rationales for the inclusion of each of the assessment measures used. The same tester administered all measures thereby reducing the effect of tester differences on the participants' scores.
Appropriate statistical analyses were completed. A reliability analysis was completed using the data for 5 of the 30 participants resulting in high reliability scores for all measures (86-93%). Overall, this study presents somewhat suggestive evidence of an inconsistent association between language and nonverbal IQ in children with language impairment.

The purpose of the study by Krassowski and Plante (1997) was to examine the stability of children's IQ scores over time in order to determine the validity of using cognitive referencing for children with SLI. This study retrospectively examined psychological test data from 75 students (male=55, female=20) in a single school district. A predetermined set of inclusion criteria was followed when selecting student files, including being tested twice on the WIPPSI, WISC-R, or WISC III. Each student's full scale, verbal scale, and performance scale (hereafter referred to as nonverbal IQ) scores on these tests were statistically analyzed.

Results of a Pearson Product-Moment Correlation found a moderately significant correlation between the testing time one and two. A within subjects ANOVA found no significant difference in scores between the two testing times; however, a significant difference was found between the three scale scores, in addition to a significant time by scale interaction. The researchers used paired comparisons to determine which IQ scale score was changing with time. Results of these comparisons found that the nonverbal IQ scale scores were significantly higher at testing time two versus testing time one. When individual nonverbal IQ scores were analyzed changes of different amounts and in either direction were observed (e.g., 15 point change=17%, 10 point change=32%, and 5 point change=64%). The results of the study revealed a significant variability in nonverbal IQ scores across time.

This study represents a 2c evidence level. The statistical analysis chosen were both reliable and valid. A weakness of this study design was that it was retrospective rather than prospective in nature. One

limitation of a retrospective study is a lack of researcher control of testing procedures. For example, having a standard length of time between the two testing times and having all children complete the same IQ measure. Additionally, the study did not have a control group of typically developing children with which to compare the findings of the experimental group. Lastly, the study used testing results from a wide range of ages (time one=3;8-11;3, time two=5;9-14;6) this could be seen as either a strength or a weakness. It could be a strength since it allows the results of the study to be generalized to a wide range of ages. It could be viewed as a weakness since the researchers did not explain whether or not their statistical analysis factored out age. Overall, the results of this study present somewhat suggestive evidence that nonverbal IQ scores vary within the same child with SLI over time.

Expert Opinion

The following two articles are examples of expert opinion without explicit critical appraisal. The level of evidence for both of these studies is level 5.

The purpose of a research note by Plante (1998) is to examine the use of test-based criteria in the diagnosis of SLI, as outlined by Stark and Tallal (1981). Plante (1998) reviews the findings of studies that include participants with SLI. Results of this review found that current research studies inconsistently used nonverbal IQ cutoff scores. Also, studies that do not use a child's nonverbal IQ score as an exclusionary criterion provide evidence that nonverbal IQ scores are typically lower in children with SLI compared to children who are typically developing. Overall, Plante found that research evidence is lacking to demonstrate that children who score below 85 on tests of nonverbal IQ (and who are not developmentally delayed) perform differently than children who score above this cutoff.

A weakness of this paper is that Plante provides an overview of the research studies, in the absence of a critical review. Overall, this article provides weak evidence that children with SLI perform differently on tests of nonverbal IQ compared with typically developing peers.

A paper by Friel-Patti (1999) provides a review of current research on SLI in order to share new information that challenges current clinical practices. Her section examining diagnostic criteria for SLI is of relevance to this analysis. The results of the research presented in this section are as follows. Identifying children with SLI by using only test scores has limitations. Clinicians should also consider how having a language impairment effects a child's everyday life. Children with nonverbal IQ scores greater than 85 are not significantly different in their language skills or response to treatment compared to children who score in the borderline range (70-84) on nonverbal IQ tests.

A weakness of this paper is that the authors did not state the selection process or criteria that was used for including research studies in this paper. Further no critical analysis of the research was completed. Overall, this article provides weak evidence that using a nonverbal IQ cutoff score of 85 as a diagnostic criterion for SLI may not accurately identify all children with SLI.

Discussion

The difficulty in identifying children with SLI and the heterogeneity present in the way that nonverbal IQ scores are used to group children with SLI makes this a challenging clinical population to study. There is increasing concern in the profession of speech-language pathology regarding the use of nonverbal IQ scores to identify children with SLI. As a result, a growing body of research exists examining the use of nonverbal intelligence scores to group children with SLI. This evidence base provides suggestive evidence that the practice of using nonverbal IQ scores to identify children with SLI may artificially skew the resultant groupings.

The results of each of the aforementioned studies supports the growing concern regarding the use of nonverbal intelligence test scores as an exclusionary criterion for SLI. Limitations in the design of each study reviewed make it so that a definitive conclusion regarding the validity of using nonverbal IQ scores, as an exclusionary criterion when identifying children with SLI, cannot be reached. This evidence base could be strengthened by future research that includes a control group so that results can be compared and contrasted with those of typically developing children. Additionally, studies that include a control group should provide a detailed explanation as to how the control group was selected and recruited, and whether any concomitant conditions are present in the group. Studies should also ensure that testers are blinded to the participants' grouping. This will reduce the bias introduced to the study results. Future research should also aim to conduct studies that are prospective rather than retrospective in design. A prospective design allows the researcher(s) more control over ensuring that the testing procedures used are standardized across study participants. Expert opinion articles should include a description about how the studies were selected for inclusion. This would allow the reader to be able to evaluate how representative of the current research the articles presented are. Additionally, expert

opinions would be of more use if they provided a critical analysis of the articles rather than just presenting the current research findings. Lastly, like many clinical studies in the field of speech-language pathology, future studies should aim to increase sample sizes. This will improve the likelihood that the findings will generalize to the greater population.

Since each study focused on a different aspect of the use of nonverbal IQ as a criterion for identifying children with SLI, it is difficult to make a broad conclusion about its use. However, the results point to the following areas for consideration. First, there may be considerable variability in the nonverbal IQ scores of children with SLI. Second, children with SLI may have lower than average nonverbal IQ. Further, nonverbal IQ scores may not be stable over time in children with SLI. Lastly, results of SLI studies may lack the ability to be compared to other studies and to generalize to the clinical population when different nonverbal intelligence tests are used.

Clinical Implications

The following recommendations for clinical practice are made based on the research evidence presented:

1) Nonverbal IQ scores can be used as a method for ruling out a general cognitive disorder; however, other applications, such as using nonverbal IQ as an indicator for service in children with SLI should be avoided or done with caution.
2) Clinicians should be mindful of the nonverbal intelligence test that was used, as different scores may be obtained depending on the test used.

Conclusions

The evidence base validates current concerns in the field of speech-language pathology regarding the use of nonverbal intelligence test scores as an exclusionary criterion for SLI. Due to the limitations present in the studies reviewed, an unequivocal conclusion cannot be reached presently. In general, the current research findings do not support the use of nonverbal IQ scores as an exclusionary criterion for identifying children with SLI.

References

Dethorne, L. S., & Watkins, R. V. (2006). Language abilities and nonverbal IQ in children with language impairment: Inconsistency across measures. *Clinical Linguistics & Phonetics*, 20(9), 641-658.

Friel-Patti, S. (1999). Specific language impairment: Continuing clinical concerns. *Topics in Language Disorders*, 20(1), 1-13.

Krassowski, E. & Plante, E. (1997). IQ variability in children with SLI: Implications for use of cognitive referencing in determining SLI. *Journal of Communication Disorders*, 30, 1-9.

Miller, C. A. & Gilbert, E. (2008). Comparison of performance on two nonverbal intelligence tests by adolescents with and without language impairment. *Journal of Communication Disorders*, 41, 358-371.

Plante, E. (1998). Criteria for SLI: The Stark and Tallal legacy and beyond. *Journal of Speech, Language, and Hearing Research*, 41, 951-957.

Rice, M. L., Tomblin, J. B., Hoffman, L., Richman, W. A. & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment: Relationships with nonverbal IQ over time. *Journal of Speech, Language, and Hearing Research*, 47, 816-834.

Stark, R. E. & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders*, 46, 114-122.

Webster, R. I., Erdos, C., Evans, K., Majnemer, A., Kehayia, E., Thordardottir, E., Evans, A. & Shevell, M. I. (2006). The clinical spectrum of developmental language impairment in school-aged children: Language, cognitive, and motor findings. *Pediatrics*, 118, e1541-e1549.