A Canadian Normative Sample for the Preschool Language Assessment Instrument

Un échantillon normatif canadien pour l'instrument d'évaluation linguistique préscolaire du langage

by • par

Elizabeth Skarakis-Doyle, PhD and William Yovetich, PhD
University of Western Ontario

Kristen Strauss, MClSc
London, Ontario

Andrea Storie, MClSc
London, Ontario

Linda Levy Fisk, MSc
High River, Alberta

Donna Torrie, MClSc
Markdale, Ontario

ABSTRACT

The *Preschool Language Assessment Instrument* (Blank, Rose, & Berlin, 1978a) is a test of preschool children's discourse abilities. Specifically, it assesses their ability to cope with the language of instruction. Blank et al.'s normative data are limited by sample size, geographic region and sociocultural background, making it difficult for Canadian speech language pathologists to use this test for identification of children with language impairment. The present study describes a normative sample collected in Ontario of 152 children between the ages of three and five years of age. Mean scores, standard deviations, standard error of measurement, and a confidence interval are provided for each of the three age groups at all four test levels. Both similarities and differences were revealed between this Ontario cohort and the normative group comprising Blank et al.'s norms. However, age group scores are more stable in the Ontario cohort due to its larger sample size making them more reliable than Blank et al.'s norms.

ABRÉGÉ

L'instrument d'évaluation linguistique préscolaire (Blank, Rose et Berlin, 1978a) est un test visant à évaluer les habiletés langagières des enfants d'âge préscolaire. Plus précisément, ce test évalue leur aptitude à négocier la langue d'enseignement. Les données normatives publiées sont limitées de par la taille de l'échantillon, la portée géographique et les antécédents socioculturels; cela fait en sorte qu'il est difficile pour les orthophonistes canadiens d'employer ce test pour dépister les enfants ayant des troubles du langage. La présente étude décrit un échantillon normatif recueilli en Ontario chez 152 enfants de trois à cinq ans. Des donneés moyennes, des écarts-types, une erreur type sur la mesure et un intervalle de précision sont fournis pour chacun de ces groupes d'âge et pour chacun des niveaux du test. Les similitudes et les différences ont été révélées entre cette cohorte ontarienne et le groupe normatif affichant les normes publiées. Cependant, les données de groupe d'âge sont plus stables chez la cohorte ontarienne vu l'échantillon plus important, ce qui les rend plus fiables que ceux des normes publiées.

KEY WORDS

assessment • language of instruction • language impairment • preschool language assessment

A critical feature of well-constructed standardized tests is an adequate normative sample. In order for a normative group to be appropriate, it must be relevant, representative and recent (Elliot & Bretzing, 1980; Gronlund & Linn, 1990). Of these criteria, representativeness in particular is typically not met in widely used standardized language tests when employed in Canada. Canadian speech-language pathologists have long laboured with unsatisfactory test norms for the language and speech tests they use. Perhaps test developers believe that the market is too small to justify the financial expense involved in developing Canadian norms. Perhaps speech-language pathologists as consumers have not made their needs known forcefully enough to developers (McCauley, 1989). Or perhaps, both developers and consumers have assumed that norms collected in the U.S. are close enough to be valid for use in Canada (Flipsen, 1993). Despite the potential reasons for the lack of Canadian normative samples, the issue of the representativeness of published norms has long been of concern to Canadian speech-language pathologists, and with good reason. Standardized tests measure more than particular language or speech abilities (Overton, 1996). The background experiences of individuals in their educational, social, and cultural environments (Salvia & Ysseldyke, 1988) are also assessed. Such experiences are not necessarily homogeneous within a geographical location such as North America, they may, in fact, vary within a given area. Differences in geographical location may result in varied experiences, values, and knowledge bases (Salvia & Ysseldyke, 1988). This fundamental issue is addressed in the *Standards for Educational and Psychological Testings* (American Psychological Association, 1985). "Norms that are presented should refer to

clearly described groups. These groups should be the ones with whom users of the test will ordinarily wish to compare the people who are tested" (p. 33). That is, the norms used should be collected from the population that the children will in fact be compared to for service eligibility or other assessment issues. In Canada, the relevant comparison is to other Canadian children not to American children. Accordingly, the development of local normative data may provide the most representative data for the comparison of individuals within a given geographical location (Anastasia, 1988).

The need for Canadian norms or even local provincial norms assumes more importance when considering that performance differences have been found in some standardized tests between Canadian and American samples. Flipsen (1993, 1998) has shown that tests in which the child's experiential base is tapped, as in vocabulary, do show differences between Canadian and American children. Further, discourse abilities by their nature are sensitive to many aspects of context including the sociocultural context. Hence, it is not unreasonable to expect differences in the discourse abilities between Canadian and American children who experience different social and cultural backgrounds. The Preschool Language Assessment Instrument (PLAI; Blank et al., 1978a) is a test of discourse in which such differences might appear.

The PLAI was created to assess a preschooler's ability to respond to the language of instruction. It does so in the context of the teacher-student exchange or discourse which is the primary context of language in the classroom (Blank, Rose, & Berlin, 1978b). It assesses one dimension of complexity in teachers' language formulations directed toward a child. That dimension is perceptual-language distance or the distance between the material being discussed (perception) and the language the teacher uses to direct analysis of the material (language). As the complexity of the formulation increases, the distance between perception and language also increases. At the same time, the demand on the child to abstract information from material available to them is increased (Blank et al., 1978b).

Four levels of abstraction in teacher formulations are assessed: These levels include: (a) Matching Perception, (b) Selective Analysis of Perception, (c) Re-ordering Perception, and (d) Reasoning about Perception. The first level, Matching Perception, emphasizes basic information that is easily available to the child from the verbal and visual material presented simultaneously with the stimulus question. The second level of formulation, Selective Analysis of Perception, purportedly requires the child to resist the attraction to the obvious characteristics of the stimuli and respond to selective features; however, all of the perceptual information the child requires to answer each question is available or predictable from the stimuli presented. The third level of teacher formulation, Re-ordering Perception, requires the child to move beyond attention to features. The child must integrate the perceptual information or actions that are obvious and manipulate his experiences to meet the verbal demands of the task. Finally, the fourth level of language used, Reasoning about Perception, represents the greatest distance between language and perception. The child must go beyond the perceptual information to reflect on implications and interpret their significance, as well as making hypotheses or predictions about possible future consequences.

Blank et al. (1978a) created the PLAI to address two objectives. The first goal was to provide a means for determining which levels of language abstraction children are capable of dealing with in the classroom. Typically the PLAI appears to be used for this descriptive purpose so that intervention can be based on its model of cognitive complexity (Moeller, Osberger, & Eccarius, 1986; vanKleek, Gillam, Hamilton, & McGrath, 1997). The second objective of the test, according to Blank et al. (1978a), was to identify children who may be "at risk" for language and academic difficulties. However, after almost 20 years, the test continues to rely on the original small experimental normative sample.

Blank et al. (1978a) conducted the preliminary standardization testing for the *PLAI* on a sample of 120 preschool children from an urban northeastern U.S. community. The sample was divided into six age groups of 20 children each, with groupings at six month interval ranging from 3;0 - 5;11. Each group had approximately equal numbers of boys and girls. Background information regarding ethnicity and parental occupation was obtained after selection. The sample was further divided by SES, half of the children in each group were from middle class homes and the other half from lower class homes. The assignment of SES was based upon where the child's daycare or nursery school was located. "Specifically, the lower-class children were drawn from 13 day care centres located in poor (inner city) neighbourhoods in New York City; the middle class children were drawn from 11 private nursery schools in more affluent sections of the city and neighbouring suburban areas" (Blank et al., 1978b, p. 51). The two stratification variables of age (in 6-month intervals) and SES resulted in groups of 10 children each. The 10 children per age-SES subgroup fall substantially short of the 100 participants per subgroup "gold standard" outlined by Salvia and Ysseldyke (1988). Another problem with this sample is that there is no evidence that children with communication impairments were included. As

McCauley & Swisher (1984) have convincingly argued, when children with communication impairments are not included in the study, the relevance of the norms for the purpose of identifying children with language impairments may be compromised.

Thus, the normative data provided in the test manual do not meet the requirements of adequate sample size or representativeness, particularly for use in Canada. Additionally, the language of instruction may differ among Canadian and American teachers and their preschool students, thus warranting the development of Canadian norms for this test. Local normative scores would be extremely useful in that they provide the most sensitive data upon which to base assumptions about an individual's performance (Anastasi, 1988).

Therefore, the purpose of this study was to obtain more adequate and representative normative data for the *PLAI* than are currently available for use with Canadian children. Specifically, the following questions were posed:

1. How do 3, 4, and 5 year old children in Ontario perform on the PLAI?
2. Do Canadian preschool children perform similarly on the PLAI to Blank et al.'s sample of urban northeastern U.S. children?

Table 1. Characteristics of the Ontario Sample of Preschool Children.

	3 year olds N=50	4 year olds N=50	5 year olds N=52	Total Sample N=152
Boys	30 (60%)	22 (44%)	33 (63.5%)	85 (55.9%)
Girls	20 (40%)	28 (56%)	19 (36.5%)	67 (44.1%)
SES 1	5 (10%)	15 (30%)	16 (30.8%)	36 (23.7%)
SES 2	20 (40%)	13 (26%)	18 (34.8%)	51 (33.6%)
SES 3	25 (50%)	22 (44%)	18 (34.6%)	65 (42.8%)
Normal Communication	48 (96%)	43 (86%)	44 (84.6%)	135 (88.8%)
Impaired Communication	2 (4%)	7 (14%)	8 (15.6%)	17 (11.3%)

Method

Participants

Children who participated in this study came from five preschools/daycare centres and one speech and hearing clinic in London, Ontario and a public school board in Elliot Lake (a mining community in northern Ontario). Thus, children from both rural northern Ontario and urban sourthwestern Ontario are included in the sample. From these centres, 152 children between the ages of 3 years 0 months and 5 years 11 months participated in this study; thirty-eight children were from the northern community, 20 five-year-olds and 18 four-year-olds; all others were from southwestern Ontario. All children had English as the primary language spoken at home. All but ten children passed a pure-tone hearing screening at the time of testing (American Speech-Language-Hearing Association, 1990). Those who did not pass the hearing screening were referred to their family physician, but were included in the norming sample. Demographic characteristics for the children are displayed in Table 1.

As can be seen the sample consisted of more males than females with the exception of the four year old group. There were fifty children each in the 3 and 4 year old groups and fifty-two in the 5 year old group. Age levels were divided into one year intervals: 3 years 0 months to 3 years 11 months, (M = 3;6, *Median* = 3:6) 4 years 0 months to 4 years 11 months (M = 4;6; *Median* = 4;6), and 5 years 0 months to 5 years 11 months. (M = 5;5, *Median* = 5;6). As can be seen in the virtually identical means and medians for each group, the distributions were symmetrically distributed at every age level.

Socioeconomic status was influenced by the location of the participating schools and centres. The SES data is presented as descriptive background data per Standard 4.4 of the *Standards for Educational and Psychological Testing* (APA, 1985) so that potential users of the norms may make informed decisions about the appropriateness of these norms to their use. They are not intended as a formal stratification variable. SES was determined based on parental education level. In cases where the parents' educational level differed, the parent with the highest level of education in each family was selected. The cohort was then divided into three groups: high school education or less (Group 1), one to three years of post-secondary education (Group 2), and four or more years of post-secondary education (Group 3). As can be seen in Table 1, slightly less than half the sample fell in Group 3, with approximately another third falling in Group 2. The remainder, and smallest proportion, fell in Group 1. The median SES for the northern cohort was one or high school diploma or less and for the southern cohort, three or postsecondary education -- four or more years.

In order to more closely approximate a representative sample, of the three- to five-year-old population, children with speech and language difficulties were also included. If children with such impairments were not represented in the standardization sample even the lowest scores included within the norms would indicate normal performance. It would then be difficult to determine how much

lower a score would need to be in order to indicate sub-normal performance (McCauley & Swisher, 1984). Thus, 17 children or 11% of the entire cohort (Beitchman, Nair, Clegg, & Patel, 1986; Flood, 1994) were identified by either a speech-language pathologist as having language (10 children), articulation, phonology (five children), and/or fluency (two children) problems, or had parents who expressed concern about one of these areas. The majority of these children were part of the southern cohort, coming from the speech and hearing clinic. Only one of the children with a communication impairment came from the school board represented in the northern cohort. Nine additional children, with no identified speech or language difficulties but who failed the hearing screening, were also included. The tenth child who failed the hearing screening also demonstrated an articulation impairment. Finally, one child had a reported unilateral hearing loss but no other speech or language impairment. All other children had no known history of permanent hearing impairment or of other speech and language difficulties.

Procedures

Children were tested individually in a quiet setting at their preschool, their kindergarten, or at the University of Western Ontario, Child Language Laboratory by graduate students trained to administer the *PLAI* or by a Speech Language Pathologist experienced in its administration. First, a pure tone hearing screening was conducted at 20 dB HL for 500, 1000, 2000, and 4000 Hz (ASHA, 1990). Following the hearing screening, the *PLAI* was administered according to the procedures published in the test manual (Blank et al., 1978a). The *PLAI* consists of 60 items which require either verbal or pointing responses. Responses were recorded both orthographically online and on audio tape. Breaks were permitted during testing as needed. The test protocols were scored according to the guidelines provided in the test manual. Each item received a quantitative score of 0 to 3 along with a qualitative descriptor.

Reliability. A second rater who was familiar with the PLAI scoring procedures, rescored 10% of the data collected or 15 randomly chosen protocols in order to determine inter-rater reliability. Point-by-point agreement for individual tests was calculated by dividing the number of rater agreements by the number of agreements + disagreements times 100. The inter-rater reliability was 93%.

Results

First, in order to ensure comparability across the SES groups represented in our cohort, one way ANOVAs were conducted to determine whether performance differences existed between the three groups at any test level. No significant differences were found at any of the four levels of the PLAI. Given this finding the data were combined for all SES groupings for all subsequent analyses. Next, in order to address potential difference in children from northern and southwestern Ontario, the twenty 5-year-old children from the north were compared to twenty randomly chosen children from the southwestern cohort of the same age. Student's *t* -tests revealed no significant difference at any of the four levels of the *PLAI*. Additionally, eighteen 4-year-olds from the north were compared to an equal number of randomly chosen same-age peers from southwestern cohort. There was one child with a language impairment in each of the 4-year-old groups unlike the groups in the previous comparison. Again, there were no significant differences found between these groups of children. Given that no differences were found between children from these two regions, both groups were combined in all subsequent analyses.

Table 2 displays descriptive statistics on the four levels of the PLAI for the entire cohort of children in each of the three age groups. As can be seen on Table 2, there was a developmental trend in accuracy both across and within age groups. Five-year-olds scored higher than either 4- or 3- year-olds at every test level, and 4-year-olds scored higher than 3-year-olds. Within each age group, mean scores decreased with increasing levels of the test. However, the 5-year-olds demonstrated less difference between test level scores than the younger children.

To address the question of comparability of these data with those published by the test's developers (Blank et al., 1978a), 20 children at each age level were chosen at random to compare to the sample of equal size provided in the test manual. Although Blank et al.'s norms are presented in six-month intervals (1978a), examination of the score distributions provided (page 4 of test manual) revealed no disernable difference for the two six month periods within an age group. This was true for 3-, 4-, and 5-year-olds. Thus, there is no empirical support for the 6 month interval grouping. However, there were differences across the year intervals just as revealed in the present data. Thus, for purposes of this analysis, the original norming data were collapsed into year intervals rather than maintaining the six-month intervals.

Recall that Blank et al.'s middle class sample was drawn from suburban and affluent urban private nursery schools, whereas their lower class sample was drawn from inner city government subsidized daycare programs. Thus, of the two SES groups for which data are provided by Blank et al. (1978a), the Ontario sample which came from primarily a suburban community (recall that there were no significant differences in the performance between the children from the south and northern rural community) is most

Table 2. Performance on the four levels of the PLAI of 3-, 4-, and 5-year-old children from Ontario.

Age	Level 1				Level 2				Level 3				Level 4			
	M	*SD*	SEM	95% CI	*M*	*SD*	SEM	95% CI	*M*	*SD*	SEM	95% CI	*M*	*SD*	SEM	95% CI
3YR	2.06	.46	.06	1.93-2.19	1.51	.49	.07	1.37-1.65	.93	.49	.07	.791-1.06	.715	.46	.06	.584-.846
4YR	2.44	.29	.04	2.36-2.52	2.12	.42	.06	1.99-2.23	1.7	.52	.07	1.56-1.85	1.4	.51	.07	1.34-1.63
5YR	2.59	.38	.05	2.49-2.70	2.36	.47	.06	2.23-2.49	2.15	.68	.09	1.97-2.34	2.05	.72	.1	1.85-2.25

Note. SEM= Standard error of measurement; CI= Confidence Interval

comparable to the middle class sample. One-sample *t*-tests were conducted at each test level for each age group. In order to control for possible cumulative type I error, a familywise error-rate was set at $p < .0125$. Only three comparisons were found to be statistically significant. Ontario 4-year-olds from the present cohort scored significantly lower than children of the same age in Blank et al's cohort at Level I ($t = 3.98$, d.f. 19, $p < .01$) only. Ontario 5-year-olds scored significantly higher at Levels III and IV ($t = 4.14$, d.f. 19, $p < .01$ and $t = 3.23$, d.f., 19, $p < .01$, respectively). Ontario 3-year-olds did not differ significantly at any level from the 3-year-olds in Blank et al.'s sample.

Discussion

The purpose of this was study was to develop a more relevant and representative normative sample for the PLAI than currently exists, particularly for use with Canadian children. Although not exhaustive, the data from this study provide more stable, and potentially more representative norms for use by Canadian speech-language pathologists. As in the previously published norms, developmental trends were apparent in the present data. That is, older children scored higher than younger children and all children scored more accurately on lower level items than higher level items. These findings are consistent with those reported by Blank et al. (1978a) with regard to the developmental sensitivity of the test.

When considering the actual mean scores across the age groups and test levels in the two cohorts, a number of additional similarities were revealed between the data from the present study and that from Blank et al.'s norms. For example, the youngest children in both samples did not differ at any test level, the 4-year-olds were not distinguishable from Blank et al. cohort on Levels 2 through 4, and the 5-year-olds did not differ on the lowest two test levels.

However, when nonsignificant results are obtained, the question arises as to whether there was sufficient statistical power in the sample size to detect differences. The comparisons conducted were limited to groups of 20 children based upon the sample size of Blank et al.'s norms. Power calculations conducted to determine whether the sample size was sufficient for differences to be detected revealed that there was sufficient power in this sample to detect all but minute differences. This was true for the comparisons of northern and southwestern Ontarian children as well. Thus, it is reasonable to conclude that, had there been differences between our groups they would have been detected, and in fact three cases were detected. This, of course, does not mean that the children in groups with no statistical difference are the same. They all reached approximately the same mean score, but could have done so in different ways or, potentially, despite their different educational and cultural backgrounds. It was beyond the scope of the current study to determine the influences upon childrens' performance.

Unique features of the Ontario cohort were also revealed. First, nine children failed the hearing screening but were included in our sample. Only one of these children scored substantially below the mean for similar age children, all others were either at or above the mean for their age group. Confirmation of permanent hearing loss or history of chronic middle ear infection was present for only two cases, and neither of these children fell outside the average for the group. All other cases are best considered as children who did not pass the screening at the time of participation and hence, their hearing status was unknown. No information is available regarding the hearing status of the children comprising Blank et al.'s norms.

Differences also emerged in comparisons of the two cohorts. Four-year-olds from the Ontario cohort scored significantly lower at Level 1 than the mean score in Blank et al.'s norms. This was an unexpected finding given that

Level 1 consists of the easiest items on the test and that they were the only age group to demonstrate a difference at this level. Several explanations are possible for this finding. First, the test-retest reliability and internal consistency (another measure of reliability) are poorest for Level 1 items (see test manual page 7). Thus, it is plausible that this difference may simply be a spurious finding due to measurement error. Specifically, three Level 1 test items (#1, #30, and #60) appeared particularly problematic for a substantial proportion of the 4-year-old Ontarian children. Twenty to 50% of this subgroup received scores of zero on these three items. In Item 1, the child abruptly has their attention shifted to a different aspect of the picture than they had just been prepared for, and in Items 30 and 60, children are not given prompting or a warning that they will be asked to recall what they had just seen. Items with characteristics such as these may contribute to the poor reliability of Level 1 items as a whole and hence, to the difference found in our results. Alternatively, in the random selection of 4-year-olds from the Ontario cohort, several children with outlying (i.e., extremely low) scores but reportedly normal communication were chosen as part of the comparison group. Their scores may have depressed the overall group mean resulting in a significant difference from Blank et al.'s data. Had a different sample been drawn in the randomization, or if it had been possible to run the entire group of Ontario 4-year-olds, the difference might not have emerged. In either case, it would appear that test related or procedural factors were most likely responsible for this difference.

There were also particularly robust differences between the two groups of 5-year-olds at the two highest test levels (which have substantially better reliability than Level 1 items). Ontario children in this age group scored significantly higher than children from Blank et al.'s cohort. It is beyond the scope of this study to established the reasons for the difference revealed. One might speculate as to the possibility that the educational background of the Ontarian 5-year-olds provided them with more training using language for problem solving such as is required in Levels 3 and 4. Additional research would be needed to determine this, so at present this explanation remains speculative. However, the quantitative differences revealed for 5 year-olds at specific test levels does suggest that the "close enough to be valid" assumption is faulty and cannot be made across all age groups of preschoolers for which the PLAI is designed. Thus, the adequacy of Blank et al.'s norms for use with Canadian children should not be taken for granted.

Although there were not extensive differences between the two cohorts, the norms developed in the course of this study have several advantages to offer speech-language pathologists working with preschool Canadian children, as well as those throughout North America. First, the larger sample size at each age level (i.e., 50 vs. 20) provides greater stability or reliability of scores. For a test to be valid it must be reliable. The norms developed in the present study enhance the reliability of the test. Second, the data for the Ontario normative sample are more contemporary than Blank et al.'s norms, having been collected more recently. Hence, they reflect contemporary cultural and educational backgrounds. Thirdly, confidence intervals are provided for each age group at each test level providing the clinician with more information about typical variation in average scores due simply to sampling error. All these features enhance the test's capabilities for providing accurate detection of children with discourse impairments. Finally, the addition of children with communicative impairments ensures that the low scores are reflective of impairment. However, caution must be exercised when assessing 3-year-olds for communicative impairments. In general, the PLAI is a difficult test for 3-year-olds developing normally, as evidenced by the low mean scores obtained particularly at the highest levels. Often, 3-year-olds with speech and or language difficulties do not have sufficient linguistic ability or sufficient attention span to take this test. This factor made it difficult to include adequate numbers of these children in the present sample which in turn makes determining an appropriate cutoff between normal and sub-normal performance more problematic. Thus, using these norms, the speech-language pathologist would be limited in accurately discriminating 3-year-old children who might have language learning impairments from those who fall at the lower end of the range of normal performance.

Although, the normative sample provided by the present study may be helpful to Canadian speech-language pathologists, a number of cautions are also warranted. The issue of the representativeness of even these norms for Canadian preschool children bears further discussion. The sample in the present study could best be characterized as middle-class Anglophone children. Thus, additional data for other SES groups and cultural/linguistic groups are needed to adequately represent the mosaic of Canadian society. Further, it cannot be assumed that these norms are representative of all areas of the nation. Additional data are needed from children in Quebec, the Western and the Atlantic provinces as well. Collecting additional data would continue to enhance stability of the norms and representativeness of this initial sample. Finally, as with any test norm, it is the clinician who must ultimately decide whether the norms developed in this study constitute an appropriate comparison group given her or his specific evaluation goal.

The present study has attempted to improve the normative sample available for the *PLAI* such that speech-language pathologists might use it not only as a vehicle for collaborative consultation with teachers, but also to assist in the identification of children with language impairments. It has been a first step in examining and improving the psychometric characteristics of the test. In addition, other aspects of the test's psychometric properties (e.g., its reliability) should be investigated as well . As McCauley (1989) has stated, the most serious consequence of inadequate measurement may be inappropriate or inadequate clinical management. More work such as that reported in this study is needed so that Canadian speech-language pathologists have adequate measurement tools to ensure appropriate clinical service to preschool children with language impairments.

Author Note

This research was supported by the Ontario District - Association of Chapters of the SPEBSQSA "Harmonize for Speech" Fund. We gratefully acknowledge the cooperation of the children and parents who participated in this study, as well as the efforts of the daycare centres', preschools' and school boards' staff who facilitated the collection of the data. Without this assistance, this study would not have been possible.

Submitted: November 1997
Accepted: August 1998

Please address all correspondence: Elizabeth Skarakis-Doyle, PhD, School of Communications Sciences and Disorders, University of Western Ontario, London, Ontario N6G 1H1.

References

American Psychological Association. (1985). Standards for Educational & Psychological Testing. Washington, DC: Author.

American Speech-Language-Hearing Association. (1990). Guidelines for screening for hearing impairment and middle ear disorders. *ASHA, 32,* 17-23.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: MacMillan.

Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G. (1986). Prevalence of speech and language disorders in 5 year old kindergarten children in the Ottawa-Carlton Region. *Journal of Speech and Hearing Disorders, 51,* 98-110.

Blank, M., Rose, S. A., & Berlin, L. J. (1978a). *The preschool language assessment instrument.* New York: Grune & Straton.

Blank, M., Rose, S. A., & Berlin, L. J. (1978b). *The language of learning: The preschool years.* New York: Grune & Straton.

Elliot, S., & Bretzing, B. (1980). Using and updating local norms. *Psychology in the Schools, 17,* 196-201.

Flipsen, P. (1993). Use of the ROWPVT with small-town Canadian kindergarten children. *Journal of Speech Language Pathology and Audiology, 17,* 145-148.

Flipsen, P. (1998). Assessing receptive vocabulary in small-town Canadian kindergarten children: Findings for the PPVT-R. *Journal of Speech Language Pathology and Audiology, 22,* 88-93.

Flood, K. (1994). *Prevalence rates of speech and language disorders in the preschool population in London, Ontario.* Unpublished master's thesis, University of Western Ontario, London, Ontario.

Gronlund, N. D. & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: Macmillian.

McCauley, R. J. (1989). Measurement as a dangerous activity. *Journal of Speech Language Pathology and Audiology, 13,* 29-32.

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49,* 34-42.

Moeller, M. P., Osberger, M. J., & Eccarius, M. (1986). Cognitively based strategies for use with hearing imapired students with comprehension deficits. *Topics in Language Disorders,* 6, 37-50.

Overton, T. (1996). *Assessment in Special Education* (2nd ed.). Englewood Hills, NJ: Prentice Hall.

Salvia, J., & Ysseldyke, J. E. (1988). *Assessment in special and remedial education* (4th ed.). Boston, MA: Houghton-Mifflin.

vanKleek, A., Gillam, R., Hamilton, L., & McGrath, C. (1997). The relationship between middle-class parent's book-sharing discussions and their preschooler's abstract language development. *Journal of Speech and Hearing Research,* 40, 1261-1271.