

1-1-1995

The accuracy of the Denver II Screening Test for children under two years of age

Thomas Alan Kapusnak
University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/rtds>

Repository Citation

Kapusnak, Thomas Alan, "The accuracy of the Denver II Screening Test for children under two years of age" (1995). *UNLV Retrospective Theses & Dissertations*. 459.
<http://dx.doi.org/10.25669/xesi-586k>

This Thesis is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Thesis has been accepted for inclusion in UNLV Retrospective Theses & Dissertations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

THE ACCURACY OF THE DENVER II SCREENING TEST
FOR CHILDREN UNDER TWO YEARS OF AGE

by

Thomas Alan Kapusnak

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science

in

Special Education

Department of Special Education
University of Nevada-Las Vegas
May 1995

UMI Number: 1374888

Copyright 1995 by
Kapusnak, Thomas Alan
All rights reserved.

UMI Microform 1374888
Copyright 1995, by UMI Company. All rights reserved.

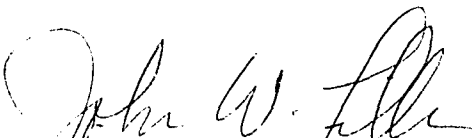
This microform edition is protected against unauthorized
copying under Title 17, United States Code.

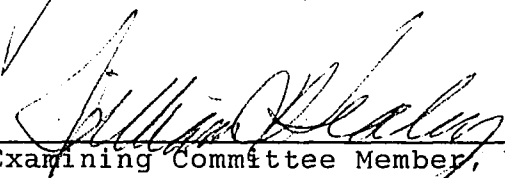
UMI
300 North Zeeb Road
Ann Arbor, MI 48103


©1995 Thomas A. Kapusnak
All Rights Reserved

The Thesis of Thomas Alan Kapusnak for the degree of
Master of Science in Special Education is approved.


Chairperson, Nasim Dil, Ph.D.


Examining Committee Member, John W. Filler, Ph.D.


Examining Committee Member, William C. Healey, Ph.D.


Graduate Faculty Representative, Lucy D. Rey, Ph.D.


Interim Dean of the Graduate College, Cheryl L. Bowles, Ed.D.

University of Nevada-Las Vegas
May 1995

ABSTRACT

**The Accuracy of the Denver II Screening Test
for Children Under Two Years of Age**

by

Thomas Alan Kapusnak

University of Nevada-Las Vegas, 1995

Program: Early Childhood Special Education

This archival study examined the accuracy of the Denver II Screening Test for children under two years of age. Accuracy was determined by comparing results from the records of eighty-two children on the Denver II Screening Test, to standards established by the American Psychological Association. These standards were: Specificity, sensitivity, and positive predictive value. Concurrent validity was established through the use of the Bayley Scales of Infant Development. The results indicated that the Denver II Screening Test was accurate for children under two years of age, especially when using the more sensitive approach in detecting children by combining questionable results and abnormal results on each test. This study recommends using the more sensitive approach to reduce the risk of missing a child in need of early intervention.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	v
CHAPTER 1 INTRODUCTION	1
Definitions and Criteria for an Accurate Screening Test	3
Statement of the Problem	5
CHAPTER 2 REVIEW OF RELATED LITERATURE	7
CHAPTER 3 METHODOLOGY AND RESULTS	11
Setting and Subjects	11
Materials	11
Procedures	12
Classification of Test Results	12
Results	13
CHAPTER 4 ANALYSIS OF RESULTS, DISCUSSION, AND CONCLUSIONS	17
Recommendations for Future Research and Conclusion	21
BIBLIOGRAPHY	24
References Cited	24
References Not Cited But Used	26

ACKNOWLEDGEMENTS

Special appreciation is expressed to my parents, who have always been supportive of my academic endeavors throughout the years. I also thank the family and friends who have supported me. I express appreciation to my committee chairperson, Dr. Nasim Dil. I also express appreciation to Dr. John Filler and Dr. William Healey for serving as committee members, and to Dr. Lucy Rey for serving as the graduate faculty representative. I acknowledge Dr. Alice Corkill and Eric Strong for their technical assistance, and the Knights of Columbus for the financial support of my graduate study. Finally, special thanks to the staff of Special Children's Clinic for their encouragement and approval of this study.

CHAPTER I

INTRODUCTION

This archival study will examine the accuracy of the Denver II Screening Test with a population of 82 children under two years of age. Previous studies (Glascoe, Byrne, Ashford, Johnson, Chang, & Strictland, 1992; Glascoe & Byrne, 1993) have measured the accuracy of the Denver II Screening Test, but had low samples of subjects under two years of age. In the past, various measures have been used to determine the accuracy and validity of both the Denver II Screening Test and the original Denver Developmental Screening Test. Developed in 1969, the Denver Developmental Screening Test (DDST) has been extensively studied to determine whether or not the DDST is an accurate screening test. The major criticism of the DDST in these studies (Applebaum, 1978; Bettenberg, 1985; Glascoe, Martin, & Humphrey, 1990; Meisels, 1989), was the low sensitivity of the DDST. The low sensitivity of the DDST was caused by the lack of referrals of children who were in need of further diagnostic testing. The language section of the DDST was determined to be an area that seemed to be missing

children with language delays. Therefore, the authors of the DDST addressed this criticism by adding twenty-two language items to increase sensitivity.

Other criticisms of the DDST were: (1) Poor predictive validity measured in longitudinal studies (Meisels, 1989), and (2) the low degree of relationship between the DDST and the Bayley Scales of Infant Development when measured by coefficients of correlation (Applebaum, 1978). The poor predictive validity of the DDST was an inappropriate criticism because the DDST, as well as the Denver II Screening Test, are not meant to be used to predict future adaptive or intellectual ability (Frankenburg, Dodds, Archer, Bresnick, Maschka, Edelman, & Shapiro, 1990). To use the DDST and the Denver II Screening Test for prediction would be ignoring the content validity of each test. Measuring predictive validity of a child's developmental level is difficult to do as developmental tests are designed to only measure a child's current level of functioning. Looking at how a child scores in the future compared to the present has little relevance when determining which children need further diagnostic testing. While screening tests will help determine which children need further diagnostic testing, the tests should not be used to predict how a child will develop in the future. Meisels (1989) cites sev-

eral longitudinal studies to show the poor predictive validity of the DDST. The weakness in doing so, however, is that the DDST was not designed to predict the future development of children.

As for measuring the degree of relationship between the DDST and a diagnostic test, a correlational coefficient should not be used. A high degree of relationship between a screening test and a diagnostic test is not as important as a screening test that meets or exceeds the standards of an accurate screening test (American Psychological Association, 1985). Screening tests and diagnostic tests are designed to serve different purposes, they won't necessarily correlate with the other. Wolery (1989) commented on the weakness of determining the accuracy of a screening test using a correlational study. "This approach has an inherent weakness, because the real issue is how well a screening test selects given students who will also score poorly on the criterion test. Thus, the real issue is not the correlation coefficient, but the "hit rate", which is determined by calculating a test's sensitivity and specificity" (p. 127).

Definitions and Criteria for an Accurate Screening Test

Screening: The application of a simple accurate method for determining which children in the population are

likely to be in need of special services in order to develop optimally. Screening procedures should not be viewed as diagnostic; they simply divide the population into those who need diagnostic work and those who are not at risk for the condition (Dumars, Duran-Flores, Foster, & Stills, 1987, p. 111).

Accuracy: Determined by comparing the results of a screening test to the standards for screening tests. These standards are: Specificity, sensitivity and positive predictive value (Glascoe, Byrne, Ashford, Johnson, Chang, & Strictland, 1992).

These standards are defined as:

Specificity: The percentage of subjects who obtain a normal result on a screening test, then obtain a normal result on a criterion test--90% is preferred.

Sensitivity: The percentage of subjects who obtain an abnormal result on a screening test, then obtain an abnormal result on a criterion test--80% is preferred.

Positive Predictive Value: The percentage of subjects who obtain an abnormal result on a criterion test out of all the subjects who obtained abnormal results on a screening test--70% is preferred.

Another factor that will be measured is a screening test's overall hit-rate. Overall hit-rate is defined as:

Overall Hit-Rate: The total number of subjects who

obtain matching results on the screening test and criterion test (Glascoe et al., 1992). A preferred rate has not been established.

Statement of the Problem

The absence of research involving the accuracy of the Denver II Screening Test for children under two years of age is the problem underlying this study. The two studies (Glascoe, Byrne, Ashford, Johnson, Chang, & Strickland, 1992; Glascoe & Byrne, 1993) involving the Denver II Screening Test contained samples with wide age range of subjects with few under the age of two years. In the first study (Glascoe et al., 1992), the ages of subjects ranged from 3 to 72 months and only 18% (19 of 104) of the subjects were under two years of age. In the second study (Glascoe & Byrne, 1993), only 15 of the 89 (17%) subjects were under two years of age, with an age range of 7 to 70 months. Glascoe and associates admit that limitations in both studies were a wide range of ages, and a limited sample of subjects under two years of age. The present study will attempt to alleviate these limitations by restricting the age range of subjects to under two years of age.

Some of the limitations of this study were: (1) A sample with a large number of clinic-referred sub-

jects, and (2) the psychologists were not "blind" to the results of the Denver II Screening Test. First, 85% of the sample was referred by outside agencies due to suspected problems. The large sample of clinic-referred subjects has been shown to "typically produce unduly favorable sensitivity" (Glascoe & Byrne, 1993, p. 370). Second, the psychologists were aware of the results on the Denver II Screening Test before administering the criterion measure. The advantages for this approach was that interrater reliability was not a concern, test results were unaffected by a subject responding inconsistently to two different examiners, and tests were given concurrently which increased the consistency of the subject's behavior. The disadvantage to this approach was that the psychologists could have biased the result on the criterion measure, since they were aware of the Denver II Screening Test result.

CHAPTER II

REVIEW OF RELATED LITERATURE

In 1990, The Denver Developmental Screening Test was revised and restandardized as the Denver II Screening Test. Glascoe (1991) noted, "Absent are studies comparing the Denver II to diagnostic instruments in order to show its sensitivity, specificity, or concurrent validity" (p. 7). Glascoe and associates later went on to conduct the two studies that measured the accuracy of the Denver II Screening Test. These studies were: the Accuracy of the Denver II in Developmental Screening (Glascoe, Byrne, Ashford, Johnson, Chang, & Strickland, 1992) and The Accuracy of Three Developmental Tests (Glascoe & Byrne, 1993). The first study was conducted in response to the need that the Denver II Screening Test "was published without evidence of its accuracy" (Glascoe et al., 1992, p. 1221). The study included 104 subjects between 3 and 72 months of age. There was a mean age of 39 months and a standard deviation of 17.1 months. The Denver II Screening Test was administered to each subject to obtain an overall result of normal, abnormal, or questionable. Subjects who were untestable were

also included in the study. The Denver II Screening Test was followed by a battery of diagnostic tests for the purpose of obtaining a criterion score. Subjects under 30 months of age were only administered the Bayley Scales of Infant Development. A summary of the results include:

- On the Denver II Screening Test, 38% (40) of the 104 subjects obtained normal results, 33% (34) obtained questionable results, 26% (27) obtained abnormal results, and 3% (3) were deemed to be untestable.
- Eighteen (17%) of the 104 subjects received an abnormal result on a diagnostic test.
- Combining test results by grouping questionable results and subjects deemed untestable with normal results, then combining questionable results and subjects deemed untestable with abnormal results--specificity was 80% and 43%; sensitivity was 56% and 83%; positive predictive value was 37% and 23%; and overall hit-rate was 76% and 50%; respectively.

The authors concluded that the Denver II Screening Test over-refers children to further diagnostic testing, causing low specificity rates.

The second study done by Glascoe and associates was designed similar to the first study. This study

(Glascoe & Byrne, 1993) included 89 subjects between 7 and 70 months of age. There was a mean age of 39.1 months and a standard deviation of 15.92 months. The Denver II Screening Test was administered to each subject to obtain a result of normal, abnormal, or questionable. Subjects deemed untestable were also included in the study. The Denver II Screening Test was followed by a battery of diagnostic tests for the purpose of obtaining a criterion reference result. Subjects under 30 months of age were only administered the Bayley Scales of Infant Development. A summary of these results include:

- On the Denver II Screening Test, 40% (36) of the 89 subjects obtained a normal result, 32% (28) obtained questionable results, 27% (24) obtained abnormal results, and 1% (1) was deemed to be untestable.
- Eighteen (20%) of the 89 subjects received an abnormal score on a diagnostic test.
- Combining test results by grouping questionable results and subjects deemed untestable with normal, then group questionable results and subjects deemed untestable with abnormal--specificity was 80% and 46%; sensitivity was 56% and 83%; positive predictive value was 42% and 28%; and the overall hit-rate was 75% and 54%; respectively.

The authors concluded that "a satisfactory relationship between Denver II scores and criterion measures could not be found" (p. 376).

It is unclear as to why the authors did not add more subjects under the age of two for the second study, even though they pointed out a limitation in their first study was a lack of subjects under the age of two. As pointed out previously, the first study had a sample of 18% (19 of 104) under the age of two, while the second study had a sample of only 17% (15 of 89) under the age of two. Also, it was difficult to ascertain why Glascoe and associates included subjects who were deemed untestable in the studies. There is no result when a subject is deemed untestable, so the subject's test should be eliminated from the study. In conclusion, both studies appeared to be critical of the Denver II Screening Test.

Consequently, the results of the literature review indicate the need for more study of the accuracy of the Denver II Screening Test. No study to date has measured the accuracy of the Denver II Screening Test exclusively with children under two years of age. The need for an accurate screening test is especially important with the increase of programs that only work with children under two years of age. Therefore, this study will indicate the effectiveness of using the Denver II Screening Test for these settings.

CHAPTER III

METHODOLOGY AND RESULTS

Setting and Subjects

The setting was a major facility designated by the State of Nevada Department of Human Resources to provide services to children birth to three years of age. The records of one hundred subjects under two years of age were randomly selected from all the children ($n = 826$) who began treatment during the 1993-1994 fiscal year. The subjects were from the Greater Las Vegas area. Of the initial 100 subjects, 18 were excluded for the following reasons: (a) incomplete test results; (b) subjects were two years of age or older; or (c) subjects were deemed untestable. The subjects' ages ranged from 3 to 23 months with a mean age of 9 months and a standard deviation of 5.3 months. The mode and median age was 7 months, and the majority (71%) of the subjects were under 1 year of age.

Materials

The Denver II Screening Test (Frankenburg, Dodds, Archer, Bresnick, Maschka, Edelman, & Shapiro, 1990)

was used. The standards established by the American Psychological Association (1985) were used to measure criterion-related evidence of validity. The criterion measure used to determine the effectiveness of the Denver II Screening Test was the Bayley Scales of Infant Development (Bayley, 1969). The Bayley Scales of Infant Development was chosen as "it is, at present, by far the best measure of infant development" (Sattler, 1988, p. 321).

Procedures

Licensed psychologists administered both the Denver II Screening Test and the Bayley Scales of Infant Development to all 826 subjects. The records of 100 subjects were randomly selected from this sample. Systematic sampling (McMillan & Schumacher, 1993) was used for this selection process. Each record was examined, and the subject's chronological age and test results were collected. If a subject was 2 years of age or older, they were eliminated from the study. Also, if the subject did not have results for both tests, they were eliminated from the study. Through this process, 18 subjects were eliminated from the study.

Classification of Test Results

The Denver II Screening Test produces an overall

result of either normal, abnormal, or questionable. A subject may also be deemed untestable. The Bayley Scales of Infant Development produces two index scores which need to be calculated to determine the result of the test. The Developmental Index score was classified as follows:

- Normal= Developmental Index (DI) of 84 and above
- Questionable= DI of 50-83
- Abnormal= DI of less than 50

Results

Results on the Denver II Screening Test and Bayley Scales of Infant Development were collected for the 82 subjects. On the Denver II Screening Test, 32% (26) of the subjects obtained normal results, 28% (23) obtained questionable results, and 40% (33) obtained abnormal results. On the Bayley Scales of Infant Development, 51% (42) of the subjects obtained a Developmental Index score which was classified as normal, 29% (24) were classified as questionable, and 20% (16) were classified as abnormal. The results on the Denver II Screening Test and the Bayley Scales of Infant Development were then computed to find the sensitivity, specificity, positive predictive value, and overall hit-rate of the Denver II Screening Test. This process was done by combining the questionable

ABNORMAL BAYLEY RESULT

DENVER II		NO	YES
	PASS	47	2
	FAIL	19	14

Sensitivity = 88%
 Specificity = 71%
 Positive Predictive Value = 42%
 Overall hit-rate = 74%

Fig. 1. Accuracy of the Denver II combining questionable and normal results.

results on each test with (1) the normal results, and (2) the abnormal results. Both sets of results were analyzed in this study.

The first analysis of results combined the questionable results with the normal results. Figure 1 presents the results from this analysis. With this approach, sensitivity was 88%--because 14 of 16 subjects obtained an abnormal result on the Denver II Screening Test, then obtained an abnormal result on the Bayley Scales of Infant Development. Specificity was 71%--because 47 of 66 subjects obtained a normal or questionable result on the Denver II Screening Test, then obtained a normal or questionable result on the Bayley Scales of Infant Development. Positive predictive value was 42%--because out of the 33 sub-

**QUESTIONABLE OR
ABNORMAL BAYLEY RESULT**

DENVER II		NO	YES
	PASS	24	2
	FAIL	18	38

Sensitivity = 95%
 Specificity = 57%
 Positive Predictive Value = 68%
 Overall hit-rate = 76%

Fig. 2. Accuracy of the Denver II combining questionable and abnormal results.

jects who obtained an abnormal result on the Denver II Screening Test, 14 subjects also obtained an abnormal result on the Bayley Scales of Infant Development. Finally, the overall hit-rate was 74%--because 61 of 82 results on the Denver II Screening Test matched the results obtained on the Bayley Scales of Infant Development.

The second analysis of results combined questionable results with abnormal results. Figure 2 presents the results from this analysis. In this case, sensitivity was 95%--because 38 of 40 subjects obtained an abnormal or questionable result on the Denver II Screening Test, then obtained an abnormal or questionable result on the Bayley Scales of Infant Development. Specificity was 57%--because 24 of 42 subjects obtained

a normal result on the Denver II Screening Test, then obtained a normal result on the Bayley Scales of Infant Development. Positive predictive value was 68%--because out of the 56 subjects who obtained an abnormal or questionable result on the Denver II Screening Test, 38 subjects also obtained an abnormal or questionable result on the Bayley Scales of Infant Development. Finally, the overall hit-rate was 76%--because 62 of 82 results on the Denver II Screening Test matched the results obtained on the Bayley Scales of Infant Development.

CHAPTER IV

ANALYSIS OF RESULTS, DISCUSSION, AND CONCLUSIONS

The results obtained from the Denver II Screening Test and the Bayley Scales of Infant Development were analyzed combining questionable results with normal results, and questionable results with abnormal results. These results were compared to the standards for an accurate screening test, as established by the American Psychological Association (1985). One of these standards, sensitivity, has a preferred rate of 80%. The Denver II Screening Test met the standard of sensitivity in both strategies of combining results. The first strategy--combining questionable results with normal results, produced a sensitivity rate of 88%. The second strategy--combining questionable results with abnormal results, produced a sensitivity rate of 95%. These results indicate that the Denver II Screening Test is very accurate in detecting children with developmental delays.

Another standard, specificity, has a preferred rate of 90%. The Denver II Screening Test was below this standard when combining questionable and normal results with a 71% rate, and below the standard for

specificity when combining questionable and abnormal results with a 57% rate. These results indicate that the Denver II Screening Test may over-refer children for further diagnostic testing.

The final standard, positive predictive value, has a preferred rate of 70%. When combining questionable results and normal results, positive predictive value was only 42%. But when combining questionable results and abnormal results, positive predictive value nearly met the preferred standard with rate of 68%. Positive predictive value is "often cited as the most important statistic for clinicians" (American Family Physician, 1992, p. 1824). Therefore, it was important that the Denver II Screening Test have a high rate of positive predictive value.

Overall hit-rate was measured as well. Looking at the strategy of combining questionable results and normal results, the overall hit rate was 74%. In the second strategy--combining questionable results with abnormal results, the overall hit-rate was 76%. While there is no "preferred" standard for overall hit-rate, these rates appear to be adequate.

The strategy of combining test results was done in order to compare the Denver II Screening Test to a criterion measure, in this case the Bayley Scales of Infant Development. The dilemma of what procedure to

follow in an actual testing situation is another issue. For example, if a child obtains a questionable result, will he or she be referred for further diagnostic testing? This issue is decided by each individual program serving children with developmental delays. In practice, some programs may send home children who obtain questionable results, while other programs would refer them for further diagnostic testing. This decision is usually made by a cost versus benefit analysis. When a program does not refer a child for further diagnostic testing after he or she receives a questionable result on a screening test, the program saves diagnostic time and money. But it risks missing a child in need of the program's services. It seems more important to reduce the risk of missing a child in need of services, than the financial aspect. However, some programs may not have the resources to finance the possibility of additional diagnostic time. This is why a screening test with the most accuracy must be a part of these programs. The Denver II Screening Test is very effective in detecting children who need services, but may do so at the expense of over-referring children.

The results of this study have indicated that the Denver II Screening Test may be more accurate with children under 2 years of age, than children 2-6 years

of age. Comparing the results in this study to the results in previous studies (Glascoe, Byrne, Ashford, Johnson, Chang, & Strickland, 1992; Glascoe & Byrne, 1993), the majority of the standards for an accurate screening test are higher by percentage. Sensitivity was higher in this study than in the previous studies. This finding is in correspondence with Glascoe and associates who reported a higher sensitivity rate when focusing on children under two years of age. As for specificity, the results were higher in this study when combining questionable results and abnormal results, but lower when combining questionable results and normal results. This finding is consistent with previous studies measuring the accuracy of the Denver II Screening Test.

In this study, positive predictive value was identical or higher by percentage when combining questionable and normal results. It was much higher when combining questionable and abnormal results. These results support the fact that the Denver II Screening Test appears to be more accurate with children under two years of age in detecting children with developmental delays. Positive predictive value is not to be confused with the predictive validity of a test. Criterion validity was measured in this study through concurrent validity. Another way to measure criterion

validity is through predictive validity. Some studies have been done for the purpose of measuring a screening test's criterion validity by using predictive validity. A flaw in using this approach is that some children may have an established or biological risk (Raver, 1991), which increases their chance of a future developmental delay. A child may obtain a normal result at present, but have a high risk of obtaining an abnormal result in the future. Thus, those using predictive validity studies with young children must consider established and biological risk.

Recommendations for Future Research and Conclusion

During the course of this study, two limitations were mentioned. Future studies may want to attempt to address these limitations which were: (1) The large number of clinic-referred subjects, and (2) the psychologists being aware of the screening test result. The first limitation can be addressed by sampling the general population. The second limitation can be addressed by the psychologists being "blind" to the results of the screening test. Other recommendations for future research include: (3) Review the preferred standard for specificity, and (4) measure the accuracy of other screening tests--focusing on the birth to three population. The third recommendation has to do

with the standard of specificity. Specificity has the highest preferred standard at 90%. It seems equally as important, if not more important, that a screening test is sensitive and does not miss a child in need of services. Yet, the standard for sensitivity is only 80%. Thus, it is suggested that the preferred standard for specificity be reviewed for adequacy. A final recommendation is to measure the accuracy of other screening tests--focusing on the birth to three population. Because of the growing need for services for the birth to three population, finding the most accurate screening test would be beneficial. This would allow children needing services to receive them, while effectively screening out those who are not in need for services.

In conclusion, the Denver II Screening Test is accurate with children under two years of age. While the Denver II Screening Test did not meet all the standards for an accurate screening test, the overall accuracy was high. As Meisels (1989) commented, "None of these statistics can be used in isolation to assess the value or effectiveness of a test, but taken together, they provide the multiple perspectives needed for evaluating a test's validity" (p. 579). Finally, because of the increased accuracy of the Denver II Screening Test for children under two years of age,

the Denver II might be preferred for programs which provide services to the birth to three population.

BIBLIOGRAPHY

References Cited

- American Family Physician (1992). Accuracy of revised Denver Developmental Screening Test. American Family Physician, 46, 1824.
- American Psychological Association. (1985). Standards for educational and psychological testing. Washington, D.C.: Author.
- Appelbaum, A.S. (1978). Validity of the revised Denver developmental screening test for referred and nonreferred samples. Psychological Reports, 43, 227-233.
- Bayley, N. (1969). Bayley Scales of Infant Development. New York: The Psychological Corporation.
- Bettenburg, A. (1985). Assessment: Instruments and procedures for assessing young children. St. Paul, MN: Minnesota Department of Education.
- Dumars, K.W., Duran-Flores, D., Foster, C., & Stills, S. (1987). Screening for developmental disabilities. In H.M. Wallace, R.F. Biehl, A.C. Oglesby, & L.T. Taft (Eds.), Handicapped children and youth. New York, NY: Human Sciences Press, Inc.
- Frankenburg, W.K., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelman, N., & Shapiro, H. (1990). Denver II. Denver, CO: Denver Developmental Materials, Inc.
- Glascoe, F.P. (1991). Developmental screening: Rationale, methods, and application. Infants and Young Children, 4, 1-10.
- Glascoe, F.P. & Byrne, K.E. (1993). The accuracy of three developmental screening tests. Journal of Early Intervention, 17, 368-379.
- Glascoe, F.P., Martin, E.D. & Humphrey, S. (1990). A

comparative review of developmental screening tests. Pediatrics, 86, 547-554.

Glascoe, F.P., Byrne, K.E., Ashford, L.G., Johnson, K.L., Chang, B. & Strickland, B. (1992). Accuracy of the Denver II in developmental screening. Pediatrics, 89, 1221-1225.

Meisels, S.J. (1989). Can developmental screening tests identify children who are developmentally at risk? Pediatrics, 83, 578-585.

McMillan, J.H. & Schumacher, S. (1993). Research in education: A conceptual introduction (3rd ed.). New York, NY: HarperCollins College Publishers.

Raver, S.A. (1991). Strategies for teaching at-risk and handicapped infants and toddlers: A trans-disciplinary approach. New York, NY: Macmillan Publishing Company.

Sattler, J.M. (1988). Assessment of children (3rd ed.). San Diego, CA: Author.

Wolery, M. (1989). Child find and screening issues. In D.B. Bailey & M. Wolery (Eds.), Assessing infants and preschoolers with handicaps. Columbus, OH: Merrill Publishing Company.

References Not Cited But Used

- Benner, S.M. (1992). Assessing young children with special needs: An ecological perspective. White Plains, NY: Longman Publishing Group.
- Bondurant-Utz, J.A. & Luciano, L.B. (1994). A practical guide to infant and preschool assessment in special education. Needham Heights, MA: Allyn and Bacon.
- Dworkin, P.H. (1989). Developmental screening--expecting the impossible? Pediatrics, 83, 619-622.
- Dworkin, P.H. (1992). Developmental screening: (Still) expecting the impossible? Pediatrics, 90, 1253-1255.
- Frankenburg, W.K. (1994). Preventing developmental delays: Is developmental screening sufficient? Pediatrics, 93, 586-593.
- Frankenburg, W.K. & Camp, B.W. (1971). The revised Denver Developmental Screening Test: Its accuracy as a screening instrument. Pediatrics, 79, 988-995.
- Frankenburg, W.K. & Dodds, J. (1969). Denver Developmental Screening Test. Denver, CO: University of Denver.
- Frankenburg, W.K., Dodds, J., Archer, P., Shapiro, H. & Bresnick, B. (1992). The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. Pediatrics, 89, 91-97.
- Haber, J.S. (1991). Early diagnoses and referral of children with developmental disabilities. American Family Physician, 43, 131-140.
- Hanson, M.J. & Lynch, E.W. (1989). Early Intervention. Austin, TX: Pro-ed.
- McCall, R.B. (1982). Issues in the early development of intelligence and its assessment. In M. Lewis & L.T. Taft (Eds.), Developmental disabilities: Theory, assessment, and intervention. Jamaica, NY: Spectrum Publications, Inc.
- Rossetti, L.M. (1990). Infant-toddler assessment: An interdisciplinary approach. Austin, TX: Pro-ed.

- Soloman, R., Clougherty, S.L., Shaffer, D., Hofkosh, D., & Edwards, M. (1994). Community-based developmental assessment sites: A new model for pediatric "child-find" activities. Infants and Young Children, 7, 67-71.
- Thurman, S.K. & Widerstrom, A.H. (1990). Infants and young children with special needs: A developmental and ecological approach (2nd ed.). Baltimore, MD: Paul H. Brooks Publishing Co.
- Witt, J.C., Elliott, S.N., Kramer, J.J., & Gresham, F.M. (1994). Assessment of children: Fundamental methods and practices. Dubuque, IA: Wm. C. Brown Communications, Inc.