

CHAPTER III: METHODOLOGY

Overview

The research questions from Chapter I established a foundation for this thesis. Question 1 focused on the accuracy of the test administration and the scoring. If flaws exist in test administration, and are recognized, then corrections can be made so that the data collected will be accurate. Question 2 focused on the amount of training time that was allotted to teaching evaluators how to properly administer the WJR-III. If poor instruction or inadequate time for instruction is the root of the problem, then the training program needs to be readdressed in order to provide appropriate training to ensure accuracy. Question 3 focused on how evaluators perceived their amount of training. The amount of training could have a direct correlation with the evaluators' comfort level of testing.

This chapter has two sections. The first section describes a survey of thirty-three of the thirty-five public schools in Onslow County, North Carolina. Not included were the Thompson Early Childhood Center for preschool children (which does not use the same assessment instruments involved in this research) and Onslow County Learning Center (an alternative placement facility).

The second section looks at actual WJR-III protocols from two schools. These protocols were scored by evaluators who were trained in Onslow County. Two different evaluators scored the protocols to determine if the evaluators came to the same conclusions.

Survey

In order to ascertain (1) the actual amount of time devoted to training evaluators to administer and interpret the WJR-III and (2) if evaluators perceived the training time as adequate, a survey was developed and sent to chairpersons of the Special Needs Departments. The survey was a single piece of paper that contained a brief summary about me and the reason for obtaining the information. I have worked in Onslow County for the past eleven years, ten in the Special Needs Department. Most of the recipients of the survey knew me personally.

The survey section had five questions which will be covered in the following paragraphs. A copy of the survey is located in Appendix A on page 53, of this thesis. The survey questions were developed with the help of Dr. Marc Sosne, a professor in the Specialty Studies Department at the Watson School of Education at the University of North Carolina Wilmington.

The survey was distributed to eighteen elementary schools, eight middle schools, and seven high schools. These are all of the public schools in Onslow County except for one alternative school and an early learning center. The survey was sent to the chairpersons of the Special Needs Departments in each school. The chairpersons were either Special Needs teachers or guidance counselors. These subjects were chosen because they have a direct impact on testing and also on placement of students in the Exceptional Needs Program.

The survey questions were developed to obtain information relevant to the research hypothesis by determining if adequate training was given so that evaluators could properly assess students. The survey questions presented were as follows:

1. Have you been trained to administer the Woodcock Johnson III assessment?
2. How much training did you receive?
3. Do you feel you had enough training to help you accurately score the assessment?
4. Who at your school gives achievement tests?
5. What other achievement tests have you administered?

Question 1 only had two response categories. Either an individual had received WJR-III training or not. Question 2 was divided into the following categories: six hours or less, seven to twelve hours, thirteen to eighteen hours, nineteen to twenty-four hours, and above twenty-five hours. Question 3 related to individuals who had actually received the training. It provided two response categories: adequate training and not received adequate training. Question 4 asked respondents to list different individuals across the county who are conducting educational assessments. Question 5 was presented to see if alternate assessments are being used to determine placement.

Woodcock-Johnson III Achievement Battery Protocol

In order to examine the accuracy of the testing, administration and interpretation of the WJR-III, elementary evaluators' scores were compared to see if there was consistency in the way two evaluators scored the same test. Data were collected from two elementary schools, grades K-5 in Onslow County, North Carolina. One of the schools is in a rural setting with an enrollment of eight hundred students. The school's population consists of 70% white, 25% African-American, and 5% other ethnic origin. Over half of the students receive free or reduced lunch, and this school is currently receiving Title I funding. The enrollment of students who are classified as Specific

Learning Disabled is twenty-two. Of those students, ten receive services for written expression.

The second school is located within the city limits of Jacksonville. The school enrollment is five hundred, seventy-five students. This school is also predominately white (78%). The African-American population is 15% and other ethnic origin makes up 7%. A very low percentage of these students receive free or reduced lunch; they come from middle to high income families. The school was redistricted a few years ago to increase the minority population but still does not qualify for Title I funding. The enrollment of Special Needs students classified as Specific Learning Disabled is eleven. Seven of these students receive services in written expression.

In both of these schools, the methods for determining eligibility for Specific Learning Disability are the same. A vision and hearing screening must be passed. The student's social history is reviewed to determine that there are no other factors such as attendance or environmental issues that affect the academic performance. Also, medical factors must be ruled out. The final component of the eligibility process is to use the student's IQ scores (WISC-III) and compare them to the standard scores on an educational evaluation (WJR-III). The Title I elementary school that is located in a rural area will be referred to as School A. The Non-Title I elementary school located within the city limits will be referred to as School B. Copies of actual WJR-III protocols were obtained from both schools.

Out of all the students that are presented to School A's Assistance Team for strategies, approximately forty students a school year are referred for Special Needs testing. Of the referrals, 12.5% are due to behavior. Most behavior referrals are students

who have constant discipline referrals. These behaviors range from aggression with peers and authority figures to verbal outbursts that disrupt the classroom setting. Other behavior issues include the student's inability to remain on task and complete assignments. The remaining 87.5% determine if a learning disability exists. These referrals come from teachers and parents.

School B's Assistance Team receives approximately thirty referrals per year for Special Needs testing. The fewer number of students is consistent due to the lower enrollment in School B. Of those referred, 10% are due to behavior issues. School B experiences the same behavior issues as School A. The remaining 90% of the referrals are to determine if the students have a learning disability that could be impacting their academic performance. These referrals are also submitted by teachers and parents.

Seven WJR-III Protocols were randomly selected from a pool of seventy students for examination. Five evaluators scored the protocols to determine consistency in the scoring. Only Subtest 11 of the Protocol was used because this subtest specifically addresses writing samples. School A had a sample size of four protocols. School B had a sample size of three protocols.

Three of the four WJR-III Protocols from School A were administered and initially scored by three different evaluators. One evaluator scored two protocols. These evaluators are certified Special Education teachers in the State of North Carolina and have been trained to administer and score the WJR-III. All three evaluators were trained within Onslow County. School B's protocols were administered and scored by the same evaluator who was trained within the county. All protocols were then reviewed and rescored by another evaluator (trained within the county) to determine if the findings on

the protocols were consistent. All evaluators used Appendix B of the WJR-III Scoring Manual to determine accurate scoring. Additional scoring guidelines are provided on pages 50 through 57, located at the back of the scoring manual. These guidelines provide more examples on grammatical and usage errors.

CHAPTER IV: RESULTS

The following pages present data derived from the results of the survey presented to Special Needs Chairpersons (See Appendix A, page 53) and seven randomly selected protocols from the WJR-III, Subtest 11. All data were collected from Onslow County School District in North Carolina.

The data are separated into two distinct sets of results. First, the survey results will compare the distribution of survey data and the responses to the survey questions. Then data from Subtest 11 of the WJR-III Protocols will be presented to compare scores by the test administrators.

Results of Survey Data

Surveys were collected from elementary, middle and high school Special Needs Chairpersons in Onslow County. Fifteen elementary, seven middle, and five high schools responded to the survey. The total number of surveys distributed per setting was: eighteen elementary, eight middle, and seven high school. Of the thirty-three questionnaires distributed in Onslow County, six were not returned (three elementary, one middle and two high school). There is no significant difference among these schools and the ones that did respond to the survey. Therefore, the response rate is 81.8%. Figure 1 (page 29) shows a comparison between the number of surveys responded to by setting, in relation to the total number distributed within the subgroup. Figure 2 (page 30) represents the percentage of the school settings that responded to the surveys. 83% of the elementary schools, 87% percent of the middle schools, and 71.4% of the high schools responded to the survey.

Question 1 asked if the chairperson of the Special Needs Department has been trained to administer the WJR-III. Of the twenty-seven chairpersons who responded to the survey, sixteen replied that they had been trained to administer the WJR-III. Eleven

chairpersons had either not been trained or were in the process, but had not completed the training. These eleven chairpersons were not administering the WJR-III to students at this time. Figure 3 (page 32) shows the comparison between the numbers of trained chairpersons to the number of untrained chairpersons.

Question 2 was directed only to those responding yes in Question 1. Of the sixteen chairpersons who had received WJR-III training, their responses to the amount of training on how to administer and score the test varied. The least amount of training received on how to administer and score the test was six hours. Three individuals rated themselves at this level. Ten individuals received ten to twelve hours of training on how to administer the WJR-III Battery. The remaining three individuals received over twenty hours of training on how to administer and score the assessment. According to Onslow County's most recent staff development manual, the WJR-III training should now last about ten hours or two days. Overall, training that has been conducted in Onslow County has ranged from six to twenty hours. One fact that must be noted was that when Onslow County presented training for less than ten hours, it was not following recent staff development guidelines. Additionally, the evaluators who received more than twenty hours received their training through college course work, not training provided by Onslow County.

The WJR-III is comprised of twenty-two subtests. By converting the amount of time to minutes and dividing that time by the number of subtests that can administered,

an approximation can be made as to how much time was spent during the training session on learning how to administer each test. The chairpersons who received six hours of training received approximately sixteen minutes of instruction on each subtest. The chairpersons who received ten to twelve hours of training received approximately thirty-two minutes of instruction on each test. The chairpersons who received over twenty hours of training received approximately fifty-four minutes of instruction on how to administer each portion of the assessment. Figure 4 (page 34) compares the total amount of instruction time received to the amount of instruction time received per subtest.

Question 3 surveyed chairpersons to determine if they felt that the amount of training they received was adequate to administer the WJR-III. Of the sixteen chairpersons who responded to the question, eleven felt they had received adequate training on administering the test. Five chairpersons responded that they did not feel adequately trained on administering the assessment. The percentage of chairpersons who felt adequately trained in administering the assessment was 69%. 31% felt they should have had more training. Table 4 (page 35) shows the comparison between the numbers

of chairpersons who felt adequately trained as opposed to those who felt they needed more training.

Figure 5 (page 36) compares the amount of instruction on administering the assessment and how the chairpersons perceived the adequacy of training. 67% of those who received instruction on administering the assessment in six hours rated their training as adequate, but 67% of those who received instruction on administering the assessment

Table 4: Perception of chairpersons' instructional training time.

	Perception of chairpersons	Percentages
Had Enough Training	11	69%
Not Enough Training	5	31%

in more than twenty hours perceived that they had not received adequate training. There is not a positive correlation between amount of training time and perceptions of adequacy of training. It may be that the method of instruction and/or the instructor of the training were ineffective and that is why the evaluators feel their training was inadequate. Of the chairpersons who received ten to twelve hours of instruction on administering the assessment, 20% perceive that they did not receive adequate training. One would expect that as the amount of training increases, the evaluator should feel he/she had more adequate training. No correlation was shown between the amount of training and the perceived adequacy of the training, according to the survey. One interesting outcome is that those who received the least amount of instruction felt they had adequate training. With only six hours of training, only the high points of testing could be covered and examiners may not be given the opportunity to learn all the complexities of the test. Possibly, the group that received ten to twelve hours of instruction had the opportunity to realize just how involved the instrument can be, and therefore, felt they needed more training.

Question 4 surveyed chairpersons to determine who is administering the WJR-III at each school. It is a school-based decision as to who administers the achievement tests. Figure 6 (page 38) presents a pie graph of who is administering the assessment.

Nine schools rely solely on the guidance counselor to administer the assessment. One school uses both the counselor and a Special Needs teacher. Four schools have the

counselor, the Special Needs teachers, and diagnosticians to administer tests. Twelve schools reported that only the Special Needs teachers are administering the assessments.

Two schools reported that no one at their school could administer the WJR-III. This last piece of data is extremely important because assessments must be completed within a timeline in order to remain in compliance for initial placement for students.

Question 5 requested a list of other assessments that are being used in order to identify disabilities. Even though the WJR-III is the most widely-used educational achievement assessment in Onslow County, other assessments are available and may provide additional information which may be a more accurate measure of a student's achievement and/or ability. Table 5 (page 40) presents a list of the other assessments that are being used in Onslow County and the number of evaluators who are administering these assessments.

Results of Protocol Data

From School A, four students were randomly chosen to whom the WJR-III Writing Sample Subtest was administered. These students will be identified as Student 1 through Student 4. Evaluator 1 administered the test and those results were part of the determining factors which led to non-eligibility placement for the students. Evaluator 2 was given the protocols at a later time and asked to score them so a comparison could be made. Evaluator 2 was not given the results presented by Evaluator 1. It is important to note that Evaluator 1, from the survey conducted, felt that he/she had not received adequate training in order to administer the test. Evaluator 2 felt confident that he/she had received adequate training.

Student 1 was administered the first twelve questions. On prompt number six, the initial evaluator gave a score of two (full credit) for the sentence. The second evaluator scored the prompt as a zero. "In order to receive full credit, the sentence must be legible

Table 5: A list of alternate assessments being utilized within Onslow County, North Carolina and the number of evaluators using these assessments.

Assessments	Number of evaluators using assessment
Stanford Binet	1
Peabody Individual Achievement Test	9
Peabody Picture Vocabulary Test	2
Key Math	3
Test of Early Reading Ability	6
Test of Early Math Ability	6

Test of Early Written Language	2
Brigance	3
Wechsler Individual Achievement Test	1
Test of Written Language	4
Oral Written Language Survey	1
Kaufman Test of Educational Achievement	1

enough so that any average adult can read the response without knowing the item content” (Woodcock, et al., 2001, p. 103). The sentence looked like the following; “the bod is yisling” and was intended to say “The bird is whistling.” The initial evaluator may have been biased on scoring because he had previous knowledge of this student. This student’s age equivalent score should be reduced from second grade-second month to first grade-ninth month. This reduction is significant because the student is actually several grade levels higher in school. The initial evaluator and the second evaluator agreed on the scoring of the other items on Student 1’s protocol.

Student 2 and Student 3 were initially administered the Woodcock-Johnson III by the same evaluator. Student 2 was administered test items one through twelve. The

initial evaluator and the second evaluator disagreed on the scoring of samples four-six and samples eight-twelve. The initial evaluator gave full credit for three samples that the second evaluator gave a “no score.” Again the student is not writing a complete sentence or the sentence is illegible for any other reader. The second evaluator reduced the score from first grade-seventh month to Kindergarten-first month. Student 3 was administered samples seven-eighteen. The initial evaluator and the second evaluator had a discrepancy on one sample. The student was given credit for a sentence fragment that should have received a “no score.”

Student 4 was administered questions one through ten. The initial evaluator scored the protocol with six errors. The second evaluator’s scoring was consistent with the first.

The total number of sample items administered was forty-eight. The initial evaluator and the second evaluator disagreed on the scoring of twelve items. Figure 7 (page 43) presents the raw scores from both evaluators for each student from School A. The initial evaluator scored each protocol higher than the second evaluator. The summary of protocol data from School A is that a slight discrepancy was found in Student 1 and 3. Student 1 showed only an 80% agreement and Student 3 showed a 92% agreement between scores. This is within the acceptable limits of 80%. There was no discrepancy with Student 4. Student 2, however, had a 55% agreement between scores. This is not an acceptable level of agreement; therefore, the scores from this test could not be considered reliable.

The WJR-III Subtest 11 protocols of the students from School B that were randomly selected were initially evaluated by the same evaluator. The results from this

evaluator determined eligibility. Evaluator 2 was given the protocols at a later time to provide a comparison of scores. Both evaluators were also given the survey and reported that they felt they had received adequate training.

Student 5 was administered sample items seven-eighteen. One discrepancy was noted between the initial evaluator and the second evaluator. The initial evaluator gave a score of one for a test item that the second evaluator gave a no score. The sentence did not fully address the prompt and should not have received a score.

Student 6 was administered sample items seven-eighteen. Two discrepancies were noted between the evaluators. The second evaluator gave both samples a no score because Student 6 did not properly address the prompt. Student 6 was to write only one sentence, but wrote two to address the prompt. The initial evaluator scored the responses as a one.

Student 7 was also administered sample items seven-eighteen. Four

discrepancies were discovered between the evaluators. The initial evaluator again gave full credit for a response that was written as two sentences instead of one. The initial evaluator and the second evaluator for School B had discrepancies on all chosen protocols. Thirty-six sample items were administered to School B. The initial evaluator and the second evaluator scored seven sample items differently. Figure 8 (page 46) presents the raw scores from both evaluators for the students at School B. While there is still a discrepancy between evaluators at School B, the discrepancy is much smaller. Student 7 had the lowest agreement or a 24% discrepancy. 76% is an unacceptable agreement. Student 5 had an 88% agreement and Student 6 had an 87% agreement which were within acceptable limits.

In summary, inaccuracy was shown in the administration and interpretation of WJR-III scores. Of the seven students at the two schools, two students received unacceptable scoring which impacted their eligibility to receive Special Education services. This is slightly less than one-third (29%) of the evaluators' scores that were inconsistent. The evaluator of Student 2 felt he had not received adequate training on administering the test, and the results prove that the non-placement was due to incompetent test interpretation. Therefore, in this case, there is a direct correlation between perceptions of adequate training and appropriate test administration and interpretation.

This chapter presents and analyzes the data from surveys completed to all of the Special Needs chairpersons within the Onslow County, North Carolina School District. Data from each question were analyzed and reported. Data were also compared between questions to determine correlations in the results. Data were also presented from WJR-III

Subtest 11 writing samples to show that test interpretation of this subtest was open to human error.

In the next chapter, specific references will be made to the results of this study, and connections will be made to the research that was proposed in Chapter I that support the findings. Additionally, strengths and weaknesses of the study will be pointed out and explained. Lastly, a need for further research will be discussed.

Figure 8: Discrepancies between evaluators' scores on same student protocols for School B.