

TEACHER PERCEPTIONS OF DIAGNOSTIC INFORMATION PROVIDED BY
STANDARDS-BASED TESTING

By

KELLI ALISE TAYLOR

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN EDUCATION

UNIVERSITY OF FLORIDA

2006

Copyright 2006

by

Kelli Alise Taylor

TABLE OF CONTENTS

	<u>Page</u>
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
ABSTRACT.....	vi
CHAPTER	
1 INTRODUCTION.....	1
Background.....	1
Purposes of Study.....	4
2 METHODS.....	6
Participants.....	6
Measures.....	7
Procedures.....	8
3 RESULTS AND ANALYSIS.....	9
Results.....	9
Analysis.....	10
4 DISCUSSION.....	16
5 LIMITATIONS AND FURTHER RESEARCH.....	20
REFERENCE LIST.....	22
BIOGRAPHICAL SKETCH.....	23

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Percent of teachers responding on each level of scale.....	12
2	Mean responses and standard deviations.....	12
3	Percent of teachers responding at each level of the scale.....	13
4	Mean responses and standard deviations to FCAT statements.....	13
5	Reliability for subject analysis.....	14
6	Responses grouped by subject.....	14
7	ANOVA results when responses grouped by subject.....	14
8	Reliability for achievement level analysis.....	14
9	Responses grouped by student achievement level.....	14
10	ANOVA results when grouped by student achievement level.....	14
11	Group responses by teaching experience.....	15
12	Mean response differences between new and veteran teachers.....	15

LIST OF FIGURES

Figure		Page
1	Grade level taught by teachers.....	7
2	Years teaching.....	7

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts in Education

**TEACHER PERCEPTIONS OF DIAGNOSTIC INFORMATION PROVIDED BY
STANDARDS-BASED TESTING**

By

Kelli Alise Taylor

August 2006

Chair: M. David Miller
Major Department: Educational Psychology

Diagnostic feedback is an important feature of some assessments. It provides help in identifying specific strengths and weaknesses. This study analyzes teachers' perceptions of the diagnostic information provided by standards-based assessments. The study specifically looks at the perceptions of the diagnostic information provided by the state of Florida's mandated Florida Comprehensive Achievement Test. Overall, teachers seemed to disagree with the idea that the FCAT provides useful diagnostic information. Analyses were done to test for differences between perceptions based on subject area (math, reading, science, writing), student achievement-level, and teaching experience.

CHAPTER 1 INTRODUCTION

Background

The No Child Left Behind Act of 2001, signed into law by President Bush in January of 2002, emphasizes increased accountability for educational results. According to the United States Department of Education, the “NCLB Act will strengthen Title I accountability by requiring states to implement statewide accountability systems covering all public schools and students. These systems must be based on challenging state standards in reading and mathematics, annual testing for all students in grades 3-8, and annual statewide progress objectives ensuring that all groups of students reach proficiency within 12 years.” In many states this has led to increased standard-based testing or increased emphasis on standards-based testing already in place.

In Florida, the Florida Comprehensive Achievement Test (FCAT) is the test used to satisfy the testing required by the federal law. The test is annually administered to all public school students in grades 3-10 in the state of Florida.

The FCAT was designed to measure “student achievement of the benchmarks contained in Florida’s Sunshine State Standards, which were developed with the goal of providing all students with an education based on high expectations” (Florida Department of Education, FCAT handbook 3). According to the FCAT handbook the test “provides feedback and accountability indicators to Florida educators, policy makers, students and other citizens”

(3). So, according to the FCAT handbook, the test is designed not only to provide an accountability measure, but also to provide feedback on student achievement.

The state of Florida uses the FCAT for four main purposes. One of the uses is as a measure of Annual Yearly Progress (AYP). AYP is a requirement of the No Child Left Behind Act of 2001. The act requires that each state establish a definition of AYP to determine yearly achievement of the district and school. AYP “measures the progress of all public schools and school districts toward enabling all students to meet the state’s academic achievement standards. AYP measurements target the performance and participation of various subgroups based on race or ethnicity, socioeconomic status, disability, and English proficiency” (Florida Department of Education, Fact sheet 1). In order for a school or district to make AYP 95% of all students and subgroups must participate in the state assessment program. The state also must set an annual objective of the percentage of students who will be proficient in math and reading. The annual objectives should lead to 100 percent of students proficient by the 2013-14 school year.

The other three uses of the test are requirements of the state of Florida. The state of Florida’s A+ Plan uses student achievement data from the FCAT to determine school performance grades. These grades are intended to “communicate to the public how well a school is performing relative to state standards” (Did 1). The plan was built upon the principles that “each student should gain a year’s worth of knowledge in a year’s time in a Florida public school” and that “no child should be left behind.”

School grades for the 2004-2005 school year used a point system which awarded a point for each percent of students who score high on the FCAT and/or make annual learning gains.

According to the Governor's Office the "heart and soul of the A+ Plan is to identify schools where student are not making appropriate annual learning gains, so a new comprehensive school improvement plan and unprecedented help and resources can be provided."

The final two uses of the FCAT involve student accountability. The state requires that third grade students who score at Level 1 (out of 5 levels) in reading on the FCAT must be retained. In addition to the third grade requirement, students must pass the reading and mathematic Sunshine State Standards portion of the grade 10 FCAT in order to receive a standard high school diploma from a public school.

The most important consideration in developing and evaluating an assessment is validity. According to the *Standards for Educational and Psychological Testing* (1999), "validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of test" (9). Validity is based on the specific inferences and uses of the test results. That is, the validity is based on the uses of the assessment, not the assessment itself. When evaluating test results it is extremely important to recognize what the test was designed to measure and what the results were intended to express. The validity of an assessment's uses is limited in the development process by the intended uses and interpretations and how the test is developed to match those uses and interpretations.

According to the FCAT handbook, the FCAT is a test designed to measure a student's achievement of Florida's Sunshine State Standards and in turn provide accountability measures to educators, policy makers, students, and other citizens. The handbook also claims that it provides feedback to educators, policy makers, and students.

The state of Florida's four uses of the FCAT results, listed and described above, all focus on accountability. While there is plenty of evidence that the FCAT is being used as an accountability measure, there is no evidence of uses regarding feedback.

LeMahieu and Wallace (1986) point out in their article, *Up Against the Wall: Psychometrics Meets Praxis*, that there are “two distinct purposes for test information: a diagnostic purpose and an evaluative purpose” (13). LeMahieu and Wallace refer to diagnostic testing as an “assessment of the strengths and weaknesses of individuals or programs expressly for the purpose of informing decisions about the curricular program that is delivered” (13). The “evaluative function of testing involves inferences about the value or worth of some component of, or participant in, the education system” (13). This may be at the school district, personnel or even student level.

Diagnostic feedback is an important feature of some assessments which can provide help in making instructional decisions. Useful diagnostic information often requires an assessment with more items than a typical standardized achievement test may contain for purposes of reliability at the skill level. While the achievement test may indicate overall student performance, it may be insufficient for identifying specific strengths and weaknesses without increased length for each specific diagnostic unit.

Purposes of Study

The goal of this investigation was to attain and analyze teachers' perceptions of diagnostic testing. The study focused on the lack of diagnostic information gained from Florida's state mandated FCAT test. The test provides the state with an accountability measure for students, teachers, schools, and districts, but there are some questions as to whether or not it provides adequate diagnostic information on the individual student to

teachers. Diagnostic information allows teachers the opportunity to see student strengths and weaknesses. With this information teachers have the opportunity to provide students with the instruction needed, and thus contribute to student learning.

The test not only investigated whether teachers felt that the state mandated test provided diagnostic information, but also investigated whether teachers felt it was even possible for such a test to provide this useful information.

The study looked to determine how useful the FCAT was to guide instruction, and how it could be made more useful to the teachers educating students.

CHAPTER 2 METHODS

Participants

This study was conducted using elementary school teachers from a county in north central Florida. The county has an estimated population of 191,000 people. Over 50 percent of the teachers in the county's school district have achieved an advanced degree, which is one of the highest percentages in the state of Florida.

Surveys were sent to 423 school teachers. This included all elementary classroom teachers teaching at one of 17 public elementary schools in the county's largest city. Using a list of mailing addresses provided by the county school board, all classroom teachers teaching at one of the 17 public elementary schools were selected as potential participants.

Potential participants received a two page survey and an informed consent letter detailing the purpose and use of the survey. Those choosing to participate completed and anonymously returned the survey within an 18 day period. Because responses were submitted anonymously there was no chance for follow up with those teachers not returning the survey.

A total of 71 elementary teachers volunteered to participate. Participants included teachers from all six elementary grade levels, including kindergarten, first, second, third, fourth, and fifth grade. Participants teaching experience ranged from one to 40 years.

Figures 1 and 2 show number of teachers responding from each grade level and the amount of teaching experience of the participants, respectively.

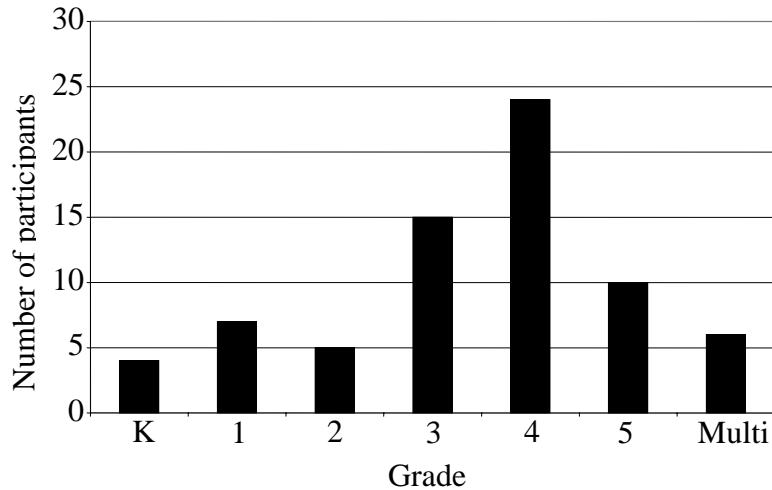


Figure 1. Grade level taught by teachers

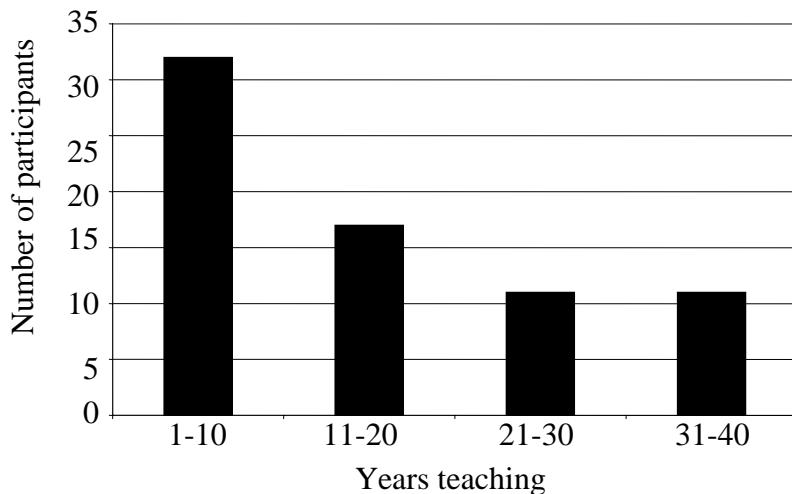


Figure 2. Years teaching

Measures

An investigator designed survey was used to collect teachers' perceptions of diagnostic testing. The survey's questions were focused around three main questions: (a) can diagnostic feedback be gained from state mandated tests? (b) does FCAT provide

diagnostic feedback? and (c) where does most of the diagnostic feedback you receive come from?

Procedures

The study was conducted using elementary school teachers. Elementary school teachers, unlike secondary teachers, have the opportunity to teach and observe students in multiple subject areas. Elementary school classrooms also tend to have a high diversity of student achievement levels within a classroom. Surveying elementary school teachers allowed the opportunity to see whether teachers perceptions of diagnostic testing varied according to student achievement level, as well as whether perceptions varied based on the subject the student was being tested on.

From a list of all elementary teachers provided by the county school board office, surveys were sent to all classroom teachers teaching at one of the elementary schools in the county's largest city. This included teachers teaching kindergarten, first, second, third, fourth, or fifth grade, as well as gifted and Exceptional Student Education (ESE) teachers. Media specialists and teachers teaching art, physical education, and music were excluded from the study.

The survey consisted of seventeen close-ended likert scale items. Participants responded to items on a five point scale ranging from strongly agree to strongly disagree. The survey also contained five open-ended items asking for information about teaching background, student testing, and general thoughts on diagnostic testing.

CHAPTER 3 RESULTS AND ANALYSIS

Results

Teachers responded to statements about the diagnostic information provided by testing on a 5 point scale. The scale ranged from 1=strongly agree to 5=strongly disagree. The tables below show the descriptive information from the data collected. Table 1 shows how teachers responded to statements regarding whether or not diagnostic information can be gained from classroom tests and whether or not diagnostic information can be gained from state mandated standardized tests. Responses to statements concerning where most of the diagnostic information teachers receive comes from are also summarized in Table 1. The mean responses and standard deviations for these statements are shown in Table 2.

Participants also responded to statements specifically about the diagnostic information provided by Florida's state mandated FCAT test. The FCAT is composed of several separate subject tests. It is composed of FCAT Math, FCAT Reading, FCAT Writing and FCAT Science. To get a more in depth look at teachers' perceptions of the diagnostic information provided, statements regarding the FCAT were broken down into subject of the test (math, reading, writing, science) as well as student achievement level (low, average, high). Table 3 shows the percent of participants that responded at each level of the 5 point scale. Table 4 shows the mean response and standard deviation for each statement.

Analysis

The descriptive statistics show that overall teachers do not agree that the FCAT provides useful feedback. To supplement this finding, further studies were conducted using the data collected to explore what factors might have an effect on the diagnostic information provided. One factor of interest was whether or not the subject of the test made a difference in teachers' perception of diagnostic information provided. Teachers' responses were averaged over achievement level, to look at overall view on diagnostic information provided by each subject test (math, reading, writing, and science). The scale reliabilities for the items grouped by subject are shown in Table 5.

A one way analysis of variance with repeated measures was done to test for mean differences between subjects. Table 6 below shows the descriptive information when responses are grouped by subject. Table 7 shows the analysis of variance results using the Huynh-Feldt correction. The p-value of 0.46 indicates that there are no significant differences between the subjects (math, reading, writing and science) and whether or not it provides diagnostic information.

A second factor that was examined was student achievement level. Teachers' responses were also averaged over subject test, to look at overall view on diagnostic information provided for low, average, and high achieving students. The scale reliabilities for the items regrouped by student achievement level are shown in Table 8.

An analysis of variance with repeated measures was done to test for any mean differences between student achievement levels. Table 9 shows the descriptive information when responses are grouped by subject. Table 10 shows the results of the analysis of variance using the Huynh-Feldt correction. The p-value of 0.08 indicates that

there are no significant differences between the mean responses for the amount of diagnostic information provided for low achieving, average achieving, and high achieving students.

The final factor that was considered in this study was the experience level of the participants. The experience level of the teachers participating in the study ranged from first year teachers to 40 year veterans. The first administration of the FCAT was in January of 1998. Teachers who have been teaching for less than 9 years have never taught without the FCAT in the school system. A test was run to test for differences in the perceptions of teachers with experience previous to the FCAT and teachers teaching since the FCAT was introduced. In order to see if there were any differences between teachers who were teaching before the introduction of the FCAT and teachers who began teaching after the introduction in 1998 an analysis of variance was conducted.

Participants were divided into two groups based on number of years teaching: new teachers (teachers teaching less than 9 years) and veterans (teachers teaching 9 or more years). The group of new teachers included 25 teachers, and the veteran teacher group had 42 teachers. For this test, only data collected from the survey's 12 questions that directly related to diagnostic information provided by the FCAT were analyzed. The results are summarized in the Tables 11 and 12. The p-value of less than .001 indicates that there is a significant difference between the mean responses of teachers classified as new and the mean responses of teachers classified as veterans.

Table 1. Percent of teachers responding on each level of scale.

1= strongly agree 5=strongly disagree.

	Level of Agreement				
	1	2	3	4	5 omitted
Diagnostic information can be gained from					
classroom tests	52.1	26.8	9.9	7	4.2
state mandated tests	14.1	26.8	33.8	16.9	8.5
Most diagnostic information I receive is from					
Teacher made tests	31	31	14.1	19.7	4.2
standardized tests	1.4	18.3	31	21.1	28.2
Overall I am satisfied with the amount of diagnostic information I receive from standardized tests	7	16.9	28.2	21.1	25.4
					1.4

Table 2. Mean Responses and Standard Deviations.

	N	Mean	Std Deviation
Diagnostic information can be gained from			
classroom tests	71	1.85	1.13
state mandated tests	71	2.79	1.15
Most diagnostic information I receive is from			
teacher made tests	71	2.35	1.23
standardized tests	71	3.56	1.13
Overall I am satisfied with the amount of diagnostic information I receive from standardized tests	70	3.41	1.25

Table 3. Percent of teachers responding at each level of the scale.

1=strongly agree 5=strongly disagree.

	Level of Agreement					
	1	2	3	4	5	Omitted
FCAT Math provides useful diagnostic information for						
low-achieving students	4.2	15.5	32.4	22.5	25.4	
average-achieving students	4.2	19.7	39.4	21.1	15.5	
high-achieving students	5.6	19.7	31	25.4	18.3	
FCAT Reading provides useful diagnostic information for						
low-achieving students	4.2	16.9	33.8	21.1	23.9	
average-achieving students	4.2	21.1	38	19.7	16.9	
high-achieving students	4.2	21.1	33.8	22.5	18.3	
FCAT Writing provides useful diagnostic information for						
low-achieving students	5.6	11.3	33.8	22.5	26.8	
average-achieving students	7	15.5	33.8	25.4	18.3	
high-achieving students	7	18.3	31	23.9	19.7	
FCAT Science provides useful diagnostic information for						
low-achieving students	5.6	8.5	32.4	23.9	25.4	4.2
average-achieving students	4.2	12.7	36.6	23.9	18.3	4.2
high-achieving students	4.2	14.1	36.6	21.1	19.7	4.2

Table 4. Mean responses and standard deviations.

	N	Mean	Std Deviation
FCAT Math provides useful diagnostic information for			
low-achieving students	71	3.49	1.16
average-achieving students	71	3.24	1.08
high-achieving students	71	3.31	1.15
FCAT Reading provides useful diagnostic information for			
low-achieving students	71	3.44	1.16
average-achieving students	71	3.24	1.10
high-achieving students	71	3.3	1.13
FCAT Writing provides useful diagnostic information for			
low-achieving students	71	3.54	1.17
average-achieving students	71	3.32	1.16
high-achieving students	71	3.31	1.19
FCAT Science provides useful diagnostic information for			
low-achieving students	68	3.57	1.15
average-achieving students	68	3.41	1.08
high-achieving students	68	3.4	1.11

Table 5. Reliability for subject analysis

Subjects	N	Number of Items	Cronbach's Alpha
Math	71	3	0.857
Reading	71	3	0.8776
Writing	71	3	0.917
Science	71	3	0.91

Table 6. Responses grouped by subject.

Subject	N	Mean	Std Deviation
Math	67	3.39	1.01
Reading	67	3.39	1.00
Writing	67	3.42	1.06
Science	67	3.47	1.03

Table 7. One-Way ANOVA with repeated measures results when responses grouped by subject.

Source of Variation	Sum of Squares	DF	Mean Square	F	Prob.
Model	0.30	2.18	0.14	0.79	0.46
Error	25.23	143.66	.18		
Total	25.53	145.84			

Table 8. Reliability for achievement level analysis

Achievement Level	N	Number of Items	Cronbach's Alpha
Low	68	4	0.9632
Average	68	4	0.9624
High	68	4	0.9626

Table 9. Responses grouped by student achievement level.

Achievement Level	N	Mean	Std Deviation
low achieving	67	3.537	1.11
average achieving	67	3.347	1.03
high achieving	67	3.358	1.08

Table 10. ANOVA results when grouped by student achievement level.

Source of Variation	Sum of Squares	DF	Mean Square	F	Prob.
Model	1.53	1.67	.914	2.69	.08
Error	37.43	110.33	.339		
Total	38.96	112			

Table 11. Group responses by teaching experience.

Groups	N	Mean	Standard Deviation
new teachers (<9 years)	12	3.55	0.14
Veteran teachers(9+ years)	12	3.33	0.10

Table 12. Mean response differences between new and veteran teachers.

Source of Variation	Sum of Squares	DF	Mean Square	F	Prob.
Between People	0.30	1	0.30	20.29	< .001
Within People	0.32	22	0.01		
Total	0.62	23			

CHAPTER 4 DISCUSSION

Concerns about standardized testing and the uses of standardized testing have been prevalent over the past decade. Questions regarding the amount of testing, type of testing, and purposes of testing are abundant among educators, politicians, and concerned parents. While there is a lot of discussion regarding the use of FCAT as an accountability measure for the state, schools, teachers and students, there is very little attention paid to the usefulness of the feedback it provides.

The results of this study confirmed that teachers feel they are not receiving sufficient diagnostic feedback from Florida's FCAT. The FCAT claims to not only provide an accountability measure, but also to provide feedback, but to this point teachers seem to be in disagreement with this claim. While the mean response to whether diagnostic information can be gained from state mandated standardized test was 2.79 (with 1 being strongly agree and 5 being strongly disagree), the mean response to whether they received diagnostic information from standardized tests was 3.56.

Beyond just looking at teachers perception on the overall diagnostic information provided by FCAT results, this study looked at several factors which might have an effect on whether or not teachers felt that diagnostic information was provided.

The first factor that was looked at was the subject of the test. Specifically, are teachers' perceptions of the diagnostic information provided different for FCAT Math, FCAT Reading, FCAT Writing, and FCAT Science? The ANOVA results for the

comparison of the mean responses about the diagnostic information provided by each subject test did not appear to be significant. That is, the subject of the test did not make a difference in the perception of the diagnostic information provided. The means for each FCAT Math, FCAT Reading, FCAT Writing, and FCAT Science were 3.39, 3.38, 3.42, and 3.47, respectively.

The same comparison was done with student achievement level. Does student achievement level make a difference in whether or not useful diagnostic feedback is provided by the FCAT? Do teachers perceive the diagnostic usefulness of the FCAT to be different for low-achieving, average-achieving, and high-achieving students? In response to these questions, the ANOVA results did not show significant differences. The diagnostic information provided was not different for different level of student achievement.

Based on the results of this study, regardless of subject matter or student achievement level, teachers seem to be in disagreement with the FCAT handbook's statement that the FCAT provides feedback. The disagreement level does not appear to be significantly different for FCAT Math, FCAT Reading, FCAT Writing, and FCAT Science. Student achievement level also does not appear to contribute a significant difference.

The one factor that had an effect on the perception of the diagnostic information provided by the FCAT was not a characteristic of the FCAT or the students taking the FCAT, but a characteristic of the participants in the study. While both new and veteran teachers disagreed with the statement that the FCAT provide useful diagnostic information, there was a significant difference in the level of disagreement. For the 12

statements specifically regarding the diagnostic information provided by the FCAT new teachers' average response was 3.55, while veteran teachers' average response was 3.33. The analysis of variance showed a significant difference in responses between teachers teaching prior to the FCAT's introduction in 1998, and teachers who began teaching after 1998. The p-value for this test was less than .001. Thus, new teachers see even less diagnostic use for the FCAT.

Whether the difference is actually related to experience with FCAT and other assessments prior to the FCAT or is related to overall teaching experience or some other factor can not be determined by the data collected in this study.

A possible reason for teachers feeling like the FCAT does not provide diagnostic information is lack of reliability. Reliability refers to the consistency of the measurement. A contributor to an assessment's reliability is the number of items. In general, the more items an assessment has, the higher the reliability. This is because the measurement is based on a more adequate sample. While the reliability of the FCAT appears to be adequate at an overall subject level, such as math, reading, and science, diagnostic information is better attained from the subparts of each subject test. These subparts typically contain only several questions. For example, the third grade FCAT Math test contains 45-50 items. This test has five subparts: Number Sense, Concepts and Operations, Measurement, Geometry and Spatial Sense, Algebraic Thinking, and Data Analysis and Probability. Together these five subparts of the FCAT Math measure 24 Sunshine State Standard benchmarks. This only equates to about two items for each benchmark if the items are distributed evenly. Teachers may feel this is too few items to provide reliable diagnostic information.

With as much time and attention teachers must dedicate to the FCAT due to the high stakes attached to the uses of the results, a push for diagnostic information is not surprising. So is there a solution to the lack of diagnostic information provided by the FCAT? If the lack of diagnostic information is an inadequate number of items to identify strengths and weaknesses, a potential solution is to add more items. This would increase the length of the test which would also increase the amount of time students are testing, so this solution is likely to receive some criticism.

Another potential solution to the lack of diagnostic information that would not increase the amount of testing time required is matrix sampling. Matrix sampling divides the set of items into different versions of a test form. A set of items is developed to cover a particular area of the curriculum and then those items are split into each of the different test forms. While this does not show specific individual strengths and weaknesses, it can provide diagnostic information for a class (Childs & Jaciw).

A disadvantage of this approach compared to the current testing is that the comparability of the student scores may decrease. Because the content is split amongst the students, students are answering different questions.

A limitation of matrix sampling is that there may not be enough items at the student level to report subscores. This lack of specific information on student achievement is already a concern for the FCAT though. While matrix sampling wouldn't necessarily provide diagnostic information for individuals, it could provide information at the classroom level to teachers.

The advantage of matrix sampling is that a broader more complete coverage of the content can be assessed without increasing the amount of testing required.

CHAPTER 5

LIMITATIONS AND FURTHER RESEARCH

There are several limitations of this study. The first limitation is the results' lack of generalizability. The participants of the study are all teachers in a north central, somewhat rural county in Florida. The perceptions of participants from this county may not be representative of all counties within the state. In more urban areas or counties with a larger population of students with limited English proficiency, teachers' perceptions of the diagnostic information may be different.

Another limitation of the study is the low response rate. Because the data was collected from an anonymously returned survey there was no opportunity for follow-up on the unreturned surveys. This led to a lower response rate.

Another limitation of the study is teachers' potential predisposition toward negative perceptions of the FCAT. Much of the pressure to perform on the FCAT falls on the teachers. Because of the pressure, teachers may have a negative view of the FCAT. This negative attitude may influence their perception of the diagnostic information provided by the FCAT.

Further research of interest might be a study looking at why the FCAT is not providing diagnostic feedback to teachers. Do the test results have the ability to provide feedback, and teachers are simply not allowed access to it? Or is the test not capable of providing the information due to lack of questions in each area?

Some follow up on why teachers teaching before FCAT's introduction differ in their perceptions of the diagnostic information provided from newer teachers might also be of interest. Is this related to having had experience with other assessments? Is it a product of having more teaching experience, or is there some other factor?

REFERENCE LIST

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing 1999*. Washington, D.C., American Educational Research Association.
- Bejar, I. (1984). Educational diagnostic assessment. *Journal of Educational Measurement* 21(2), 175-189.
- Childs, Ruth A., & Jaciw, Andrew P (2003). Matrix sampling of items in large-scale assessments. *Practical Assessment, Research & Evaluation*, 8(16).
- “Did You Know.” (2000). *MyFlorida.com*. My Florida. 5 Jun 2006
<<http://www.myflorida.com/myflorida/government/governorinitiatives/aplusplan/youKnow.html>>
- Florida Department of Education (2005). *FCAT handbook- A resource for educators*. Tallahassee: State of Florida, 2005.
- Florida Department of Education (n.d.[a]). *Fact sheet: NCLB and adequate yearly progress*. 2 May 2006 <<http://wwwfldoe.org/nclb/factsheet-ayp.pdf>>
- Florida Department of Education (n.d.[b]). *Frequently asked questions about the FCAT*. 2 May 2006 <<http://www.firn.edu/doe/sas/fcat/pdf/fcatfaq1.pdf>>
- LeMahieu, Paul G. & Wallace, Richard C. (1986). Up against the wall: Psychometrics meets praxis. *Educational Measurement: Issues and Practice* 5 (1), 12-16.
- U.S. Department of Education. (Jul 2002). *Dear colleague letter to education officials regarding implementation of No Child Left Behind and accountability, and providing guidance on adequate yearly progress*. 2 May 2006
<<http://www.ed.gov/policy/elsec/guid/secletter/020724.html>>

BIOGRAPHICAL SKETCH

Kelli Alise Taylor was born on March 18, 1982, in Panama City, Florida. She received a Bachelor of Science degree in statistics from the University of Florida in May of 2004. Kelli's parents and older brother have also all received degrees from the University of Florida. During her time in graduate school at the University of Florida she had the opportunity to serve as a teaching assistant for two courses in the College of Education.