University of New Orleans

ScholarWorks@UNO

University of New Orleans Theses and Dissertations

Dissertations and Theses

Fall 12-17-2011

Measuring early numeracy of kindergarten students in a group setting

Stacy A. Winck

University of New Orleans, sluke@bpi.nocoxmail.com

Follow this and additional works at: https://scholarworks.uno.edu/td

Part of the Educational Assessment, Evaluation, and Research Commons, and the Special Education and Teaching Commons

Recommended Citation

Winck, Stacy A., "Measuring early numeracy of kindergarten students in a group setting" (2011). *University of New Orleans Theses and Dissertations*. 1361.
https://scholarworks.uno.edu/td/1361

This Dissertation is protected by copyright and/or related rights. It has been brought to you by ScholarWorks@UNO with permission from the rights-holder(s). You are free to use this Dissertation 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 Dissertation has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. For more information, please contact scholarworks@uno.edu.

Measuring early numeracy of kindergarten students in a group setting

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Special Education and Habilitative Services

by

Stacy Winck Crane

B.A. University of New Orleans, 1992
M.C.D Louisiana State University Health Sciences Center, School of Allied Health, 1994

December, 2011

Copyright 2011, Stacy Winck Crane

Table of Contents

List of Figures

List of Tables

Abstract

Early identification practices in assessment are crucial to preventing academic failure as well as identifying students at-risk for later learning disabilities. The *PAM Early Numeracy Screening* is a set of subscales designed to measure early numeracy in kindergarten students in a group setting. Given that the existing early numeracy measures are individually administered, the purpose of the current study was to explore the psychometric properties of the *PAM Early Numeracy Screening*. Correlational analysis was the primary research design used to investigate the evidence of reliability, criterion-related validity, and construct validity of the *PAM Early Numeracy Screening*. Criterion measures included the *KeyMath-3 Diagnostic Assessment*, *Metropolitan Achievement Tests*, Eighth Edition, *Early Math Measures Study Teacher Rating of Students' Math Proficiency*, and the *Early Numeracy Indicators*. The sample consisted of 97 kindergarten students from a school district in the Greater New Orleans Metropolitan Area. Results support the *PAM Early Numeracy Screening* as a promising group administered measure of early numeracy in kindergarten. Implications for future research include investigating the internal structure of the subscales and exploring evidence of predictive validity of the subscales, specifically the Quantity Discrimination subscale to independently predict later math achievement.

key words: curriculum-based measures, early numeracy, kindergarten, group

CHAPTER 1

Introduction

Historically, mathematics performance of students in the United States lags well behind that of their peers in other developed countries. Out of 41 countries, students in the United States ranked 24th in math literacy and 26th in math problem solving (Lemke, Sen, Pahlke, Partelow, Miller, Williams, Kastberg, & Jocelyn, 2004). In international comparisons conducted by the Organization for Economic Cooperation and Development (OECD) member countries, the United States ranked 25th out of 34 countries in math literacy on the Program for International Student Assessment (Fleischman, Hopstock, Pelczar, & Shelley, 2010). Despite this disparity, in recent years there has been little or no improvement in student performance on national mathematics exams such as NAEP (National Center for Education Statistics, 2009). Student mathematics performance continues to decline.

Many students graduate from high school without the necessary math skills they need to obtain a job or attend college. In 2006, one of President Bush's main provisions in the *No Child Left Behind Act* was to develop a secondary education system that provided students with those necessary skills. However, change and development have been slow to occur. In October 2009, the National Center for Education Statistics reported results from math assessments that indicated that 82% of the students in fourth grade were at or above basic and 34% were approaching proficient. In addition, no significant change was noted in the mathematics performance of fourth grade students from 2007 to 2009 (National Center for Education Statistics, 2009). Without change to our nation's educational system, the mathematics achievement of our students will continue to decline.

Weaknesses in student math achievement begin well before high school, often in elementary school. Identification of students with math disabilities in high school or middle school is too late. Early identification coupled with appropriate intervention is crucial for the prevention of math failure. Several studies have noted that children enter kindergarten with varying levels of math knowledge and, if the differences are not addressed, weaknesses grow into deficits as students matriculate through school (Gersten, Jordon, & Flojo, 2005; Mazzocco & Thompson, 2005).

Fuchs, Compton, Fuchs, Paulsen, Bryant, and Hamlett (2005) noted that early prevention of mathematics difficulties rather than later remediation is a more promising approach to addressing math deficits. A critical first step toward the prevention of mathematics disabilities requires screening tools that produce valid and reliable data that can be used to identify children at-risk for learning difficulties. Furthermore, for students to receive early intervention, identification is imperative.

Statement of the Problem

Prevention and early identification of math deficits are in its infancy. Foegen, Jiban and Deno (2007) documented only 32 studies in math identification and progress monitoring, four of which were targeted for early mathematics. Because progress monitoring in math is in the early stages, researchers have focused on the development of technically adequate measures for second through fifth grades. Recently researchers have shifted their emphasis on the creation of early numeracy Curriculum-Based Measures to identify children in kindergarten who are at-risk for developing difficulties in mathematics learning (Chard, Clarke, Baker, Otterstedt, Braun, & Katz, 2005; Clarke & Shinn, 2004; Joyce & Wolking, 1987; Lembke & Foegen, 2009; VanDerHeyden, Witt, Naquin, & Noell, 2001). However, these measures are individually administered at a cost of student instructional time and economic impact.

Overview of the Conceptual Framework

Validity

Messick (1980) describes validity as an extensive evaluation of the information which substantiates test use and test interpretation. He clarifies that the test does not possess validity; rather, validity lies in the interpretation and use of the test scores. The score is not only a utility/ representation of the test items but also of the person taking the assessment and the conditions of the test (Cronbach, 1971). Angoff (1988) argued that validity must be viewed as a process not as a systematic procedure. All evidence, therefore, must be evaluated before test scores are deemed valid.

Curriculum-Based Measures

The principles of Curriculum-Based Measures are founded in special education research, Applied Behavior Analysis Theory, and Precision Teaching with a focus on psychometric properties. Curriculum-Based Measures were designed for the purpose of assisting teachers in measuring student progress, using repeated data collection, and in making informed data-based decisions regarding student progress and teacher effectiveness (Deno & Mirkin, 1977). As a result of the six-year federally funded project through the University of Minnesota Institute for Research on Learning Disabilities (IRLD), a general group of progress monitoring tools for reading, writing, and spelling were developed. These processes included the following criteria: "(a) the core outcome tasks on which performance should be measured; (b) the stimulus items, the measurement activities, and the scoring performance to produce technically adequate data; and (c) the decision rules to improve educational programs" (Deno, 2003, p.184). Deno and Fuchs (1987) furthered this research to include standards for technical adequacy, validity use, and the practicality of the measures.

While teachers continued to draw on curriculum-specific materials to measure student progress within special education classrooms, Fuchs and Deno (1994) developed a set of generic measurements called General Outcome Measures (GOC) for reading. Meanwhile research into math outcome measures lagged significantly behind (Foegen, Jiban, & Deno, 2007).

The initial research in the area of early numeracy Curriculum-Based Measures can be traced to Joyce and Wolking (1987) and VanDerHeyden, Witt, Naquin, and Noell, (2001). Joyce and Wolking explored the abilities of pre-school children to identify printed numbers and dots and to count backwards from 10. VanDerHeyden, et al. (2001) examined kindergarten readiness skills for reading and mathematics in a group setting. The pioneers in the field of early numeracy, however, are Clark and Shinn (2004). In their study, Clark and Shinn (2004) investigated the relationships between four individually administered experimental subscales (Oral Counting, Number Identification, Quantity Discrimination, and Missing Number) and three criterion measures (Woodcock-Johnson Applied Problems subtest; Number Knowledge Test and math Curriculum-Based Measurement first-grade computation probes). Results were promising. All four experimental subscales produced scores that were reliable for screening decisions, and all demonstrated adequate evidence of concurrent and predictive validity. Advancements continued as other researchers explored the technical features of individually administered measures of early numeracy to be used as screening tools for number sense (Chard, Clarke, Baker, Otterstedt, Braun & Katz, 2005; Clarke, Baker, Smolkowski & Chard, 2008; Martinez, Missall, Graney, Aricak & Clarke, 2005). Lembke and Foegen (2009) continued the research of Clark and Shinn (2004) using revised forms of the *Test of Early Numeracy*.

Purpose of the Study

Building on the research of Lembke and Foegen (2009) and VanDerHeyden, et al. (2001), the current study investigated the psychometric properties of the *PAM Early Numeracy Screening*, a group administered early numeracy Curriculum-Based Measure. This study was designed to fulfill the following goals: (a) to identify which subscales of the *PAM Early Numeracy Screening* produce accurate scores, (b) to identify which subscales of the *PAM Early Numeracy Screening* demonstrate a relationship with the criterion measures, and (c) to identify which subscales of the *PAM Early Numeracy Screening* demonstrate a relationship with the construct of number sense.

Rationale

Fewer than 10 studies exist in the early numeracy Curriculum-Based Measures screening literature. All studies include kindergarten and/or first grade students who were individually administered Curriculum-Based Measures of early numeracy. However, there is one study VanDerHeyden, et al. (2001) to date that reports kindergarten group administration of Curriculum-Based Measure of early numeracy. Additionally, there are only two commercially published early numeracy progress monitoring tools for kindergarten. Both products, AIMSweb and mclass Math, are individually administered.

AIMSweb Math Curriculum Based Measurement Test of Early Numeracy (AIMSweb TEN) is available through Pearson Publishing and includes four subscales: Missing Number, Number Identification, Oral Counting, and Quantity Discrimination. Each measure is individually administered for 1 minute and scoring is completed during the administration. For a class of 25 kindergarten students, a teacher may spend an average of 10 minutes per student, resulting in approximately 4 hours of testing. All measures have convincing evidence of reliability of scores and the slope, alternate –form reliability (includes 33 forms of each

subscale), validity of the performance level score as well as predictive validity of the slope of improvement, established end of the year benchmarks, disaggregated reliability and validity data, and rates of improvement (National Center on Response to Intervention, 2011). Costs may range from $39.00-$299.00 for the measurement sets or approximately $4.00 per student for the software subscription. This equates to $100.00 for a classroom of 25 kindergarten students. Additionally, AIMSweb recommends 1.5-2 hours of administration training (Aimsweb, 2011).

mClass Math contains four subscales: Missing Number, Number Identification, Oral Counting, and Quantity Discrimination. Each measure is individually administered for 1 minute, and scoring is completed during the administration. mClass Math, published through Wireless Generation, requires a hand held computer for administration and scoring. The technical adequacy of mClass Math varies according to the measure. All subscales demonstrated convincing evidence of reliability and validity of performance level scores and end of the year benchmarks. Only Number Identification and Quantity Discrimination had convincing evidence of alternate-form reliability (National Center on the Response to Intervention, 2011). For the initial year of implementation, costs may range from $1400 to $2000 for campus installation fees, handheld computers, teacher mClass Math kits, and student fees. For each subsequent year that the school uses the mClass system, the school can expect to pay $13.00 per student. mClass recommends 4-8 hours of administration training (Wireless Generation, 2011). While both measures demonstrate evidence of reliability and validity, neither mClass nor Aimsweb is a viable solution to the early prevention and identification of math deficits.

There is a need for an early numeracy Curriculum-Based Measure that demonstrates evidence of reliable and valid scores, is cost efficient for educators and school districts, and allows teachers the opportunity for increased classroom instructional time.

Research Questions

The following research questions were examined to meet the purpose of this study.

1. What is the reliability of the scores on the *PAM Early Numeracy Screening*?
2. What is the concurrent criterion-related validity of the *PAM Early Numeracy Screening*?
3. What evidence of construct validity exists for the use of the *PAM Early Numeracy Screening*?

Overview of Methodology

A quantitative validation study was used to explore the psychometric properties of the *PAM Early Numeracy Screening*. The researcher collected data to examine the relationships between the *PAM Early Numeracy Screening* and the criterion variables. For both the pilot study and current study, data were collected from a sample of kindergarten participants within a large school district in the Greater New Orleans Metropolitan Area. In the pilot study, the researcher collected data using the *Early Numeracy Indicators* (Lembke & Foegen, 2009), an individually administered early numeracy Curriculum-Based Measure, the *PAM Early Numeracy Screening-Pilot*, a modified version of *Early Numeracy Indicators*, and the Calculation and Applied Problems subtests of the *Woodcock-Johnson III Tests of Achievement* (Woodcock, McGrew & Mather, 2001, 2007). In the current study, the researcher collected data using the Number Identification subscale of the *Early Numeracy Indicators* (Lembke & Foegen, 2009), an individually administered early numeracy Curriculum-Based Measure, the *PAM Early Numeracy Screening*, a modified version of the *PAM Early Numeracy Screening-Pilot*, the Basic Concepts and Applications Composite Areas of the *Key Math-3 Diagnostic Assessment* (Connolly, 2007a), the *Metropolitan Achievement Tests*, Eighth Edition (Pearson, 2000), and the *Early Math Measures Study Teacher Rating of Student's Math Proficiency* (Lembke and Foegen, 2009).

Definition of terms

For the purpose of this study key terms are defined.

Number Sense

Number sense "reputedly constitutes an awareness, institution, recognition, knowledge, skill, ability, desire, feel, expectation, process, conceptual structure of a mental number line" (Berch, 2005, p. 333).

Howden (1989) defines number sense as having good perception and understanding of numbers and their relationships which progresses as children explore numbers in various contexts.

Early Numeracy

Early numeracy is the "basic skills and knowledge akin to a number sense" (Methe, Hojnoski, Clarke, Owens, Lilley, Politylo, White, & Marcotte, 2011, p. 200).

Curriculum-Based Measurement

Curriculum-Based Measurement (CBM) is a process that repeatedly measures students' progress using curriculum-based materials.

Quantity Discrimination

Quantity discrimination is the evaluation and interpretation of quantities using numerical symbols (Greeno, 1991).

Missing Number

Missing number is the provision in a sequential set of numerical values of the missing numerical value.

Number Identification

Number identification is the identification of a positive numerical symbol from zero to 100.

Criterion-related validity

The term criterion-related validity is defined in the following manner: "Evidence demonstrates that test scores are systemically related to one or more outcome criteria" (American Education Research Association, American Psychologist Association & National Council on Measurement in Education, 1985, p.11).

Salvia, Ysseldyke and Bolt (2007) state that "Criterion-related validity refers to the extent to which a person's performance on a criterion measure can be estimated from that person's performance on the assessment procedure being evaluated" (p. 150).

Construct validity

Messick (1980) states that construct validity "is a process of marshaling evidence to support the inference that an observed response consistency in test performance had a particular meaning, primarily by appraising the extent to which empirical relationships with other measures or the lack thereof are consistent with that meaning" (p. 1015).

Reliability

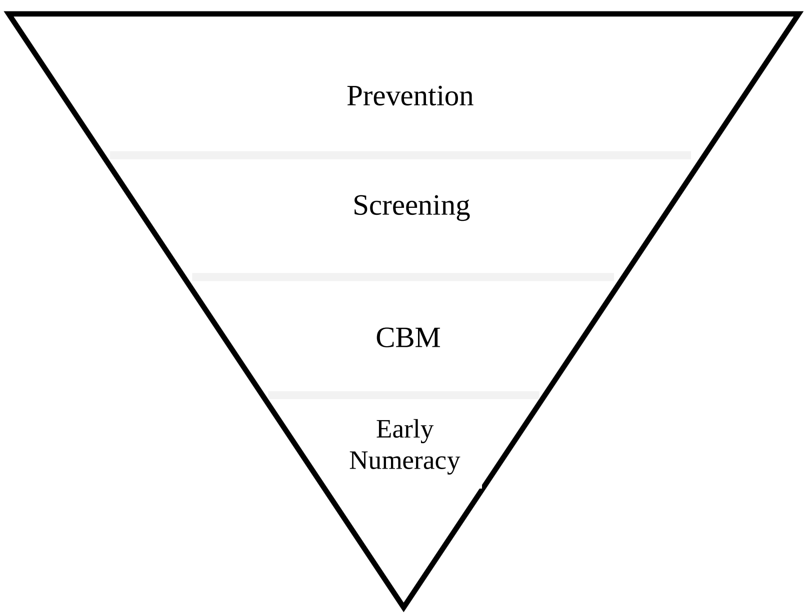
The term "reliability refers to the degree to which test scores are free from errors of measurement" (American Education Research Association, American Psychologist Association & National Council on Measurement in Education, 1985, p.19).

CHAPTER 2

Review of Related Literature

This chapter is organized by the following headings: prevention/identification, screening instruments, Curriculum-Based Measures, and early numeracy. The conceptual framework for this study is based on the assumptions that prevention will lead to early identification, that screening is an initial part of prevention and identification, that Curriculum-Based Measures are essential screening tools for reading and mathematics that produce valid and reliable scores, and that early numeracy is a crucial component to identifying students in kindergarten that have mathematics difficulties. Figure 2.1 depicts the conceptual framework of this study.

Figure 2.1 Conceptual Framework



Prevention/Identification

The first assumption is that prevention is crucial to decreasing the number of students exhibiting math difficulties. In recent years, identification of students with learning disabilities changed from a "wait to fail" model to a prevention model. In previous years, students were

identified using an IQ and achievement discrepancy model, and as a result, most students were not identified until later in elementary school. However, the 2004 reauthorization of Public Law 108-446, Individuals with Disabilities Education Improvement Act, indicates that states are no longer required to use the discrepancy model for LD identification but a Response to Intervention (RTI) model may be used. In this model, students with learning differences are identified based on their response to intervention (Wedl, 2005).

RTI is a multi-tiered problem solving model. Most educators agree on a three-tiered model for RTI. The Primary Tier is considered general education or prevention. At this level, children within the general education curriculum and students are screened using universal screening procedures. This tier consists of whole class instruction that is scientifically based and yields positive learning for the majority of the students. The Secondary Tier is a prevention level where students are provided with instruction in small groups. Students in the Secondary Tier may be experiencing some difficulty in the classroom and require additional support. The Tertiary Tier is designed to continue to provide additional instruction to students who are not responding to intervention in the Secondary Tier. It usually takes place in the general classroom in small groups or in a 1 to 1 situation. If a student has not responded to intervention at the first three tiers, then a more intensive and focus research-based approach may be warranted. Within the Fourth Tier of RTI, students are evaluated for behavioral and/or academic difficulties and are usually placed with in part-time or full-time special education environment. At each level within the model, students are monitored for progress, and students can easily amongst the tiers. The goals of RTI, therefore, are prevention of later academic or behavioral deficits and identification of students with learning disabilities who have been unresponsive to general instruction and may require individualized instruction (Fuchs & Fuchs, 2007; Horowitz, 2005).

Within the RTI model, the first level in the general education realm is prevention. Part of prevention is identification. Identification is necessary so that appropriate intervention can be employed (Kratochwill, Albers, & Shernoff, 2004). One way to identify students is through universal screening. All students are screened to identify those who might be at-risk for academic failure or behavioral difficulties and who might possibly benefit from intervention (Severson & Walker, 2002). According to Glovers and Albers (2005) to accurately identify students, screening tools must be applicable for the proposed purpose, technically adequate, and usable (i.e. cost efficient and user friendly).

Screening

The second assumption is that in order to screen students accurately, screening technically adequate instruments are essential for identifying students with math difficulties. Technical characteristics of screening instruments should include acceptable norms, reliability coefficients (internal consistency, test-retest, inter-rater), and evidence of validity (criterion-related, content and construct). Depending on the situation, school, district, or national norms may be appropriate (Glovers & Albers, 2005). Internal consistency estimates are essential for screening instruments, because they provide information on whether the items on an assessment are measuring the construct (Salvia, Ysseldyke, & Bolt, 2007). When screening instruments have multiple forms, test-retest reliability coefficients as well as the mean and standard deviation of the subtest, are important. Test-retest reliability coefficients provide information on student performance on an individual measure. In addition, inter-rater reliability coefficients are crucial especially when multi-evaluators are making judgments in regards to a student’s responses (AERA, APA, & NCME, 1999).

The reliability of scores produced by an instrument are important as well as the establishment of validity evidence of the instrument. Messick (1995) defines validity as "an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of tests scores or other modes of assessment" (p.741). AERA et al. (1999) describes five sources of evidence to be gathered for validity: test content, consequences of testing, convergent and discriminant power, internal structure, and the relationship of the assessment with other assessments. Evidence from these sources can be further categorized into content validity, construct validity and criterion-related validity.

Historically, the validity of an instrument has been measured through correlational analyses in which the scores of the proposed measure are correlated with a criterion measure. However, researchers argued that the reliability and validity of the criterion measure must also be examined, and evaluators must be careful when interpreting the results of the correlations. Furthermore, researchers believed that the operational definition of the attributes measured by the criterion need to be investigated and more than one measure of validity evidence was necessary. In 1954, APA published technical recommendations for psychological tests which are the bases for the current recommendation published by the APA Committee on Test Standards. It was during this early research that the idea of construct validity was created. The concept of construct validity began to grow and to include the idea of nomological network first defined in 1955 by Cronbach and Meehl. As the research grew to support the idea of construct validity, the APA Committee on Test Standards began to revise the standards to reflect the idea that validity was inferred and judged based on the existing evidence which may include test scores. These beliefs led to researchers stating given that all existing forms of validity (predictive, construct,

concurrent and content) were in principle part of construct validity. This idea fostered the belief that validity is a single concept under which evidence for all types of validity exists. As a result, APA et al. (1985; 1999) published the *Standards for Education and Psychological Testing* which included the idea that several sources of evidence are necessary to infer if sufficient evidence of validity exists for a measure (Sireci, 2009).

Salvia, Ysseldyke, and Bolt (2007) suggest that content validity is an important source of evidence for many educational and psychological tests, and they defined it as the degree that the items on the test represent the content domain to be measured. To examine content validity, a clear definition of the measured domain is necessary. Furthermore, when measuring the content validity of a screening instrument, Glover and Albers (2007) suggest three types of indicators to measure: item-discrimination coefficients, item difficulty indices, and differential item functioning. As a result when designing an instrument, the purpose of the assessment must be considered as well as the content domain that the test is proposed to represent (Salvia, Ysseldyke, & Bolt (2007).

An additional source of validity is construct validity. Salvia, Ysseldyke, and Bolt (2007) describe construct validity as the degree to which a test measures the theoretical trait it purports to measure. Messick (1980) states that construct validity focuses on two relationships for the test: the relationship between the test and the different procedures used to measure the same construct and the relationship between the specific construct and other constructs which are related to the specific construct on theoretical grounds. AERA et. al (1999) suggest one way to measure construct validity for a screening instrument is to examine the relationship between a similar instrument and an instrument that measures a different skill. If the there is a strong correlation between the two instruments that purports to measure a similar construct and a weak correlation

between the instrument that claims to measure a different skill, then one can infer that evidence of construct validation exists.

Another source of the validity of a test is criterion-related validity which is the degree to which an individual's score on the assessment being validated can predict the individual's score on the criterion measure. The validity of the criterion measure is critical; the measure must be a valid instrument to provide evidence of validity for another instrument. Furthermore, statistical analyses should be provided for the criterion measure as well as the generalizability of the information to other individuals and criterion measures (Salvia, Ysseldyke, & Bolt, 2007). Two types of criterion validity exist: predictive and concurrent validity. Predictive validity refers to the degree to which the assessment being validated can predict a future score, and concurrent validity refers to the degree to which the assessment can predict the score when both instruments are given at the same time. While Glover and Albers (2007) state predictive validity is the central indicator of the technical adequacy of a screening instrument, they also consider concurrent validity to be an important indicator of a screening instrument. It is essential for the screening instrument to discriminate between students who will and students who will not exhibit academic or behavior difficulties.

While it is imperative that a screening instrument demonstrate technical adequacy, the feasibility and usability of an instrument must be explored. Educators must consider the cost of the instrument and its practicality. Additionally, the personnel resources required for training, administering, and scoring the screening need to be discussed amongst school personnel. Furthermore, it is important to determine if accommodations can be provided to the targeted population and if norms are available to represent the population, as well as guidelines for administration and scoring the assessment (AERA et al., 1999). The most important outcome of

screening is that the information gleaned from the screening is useful to educators so that instruction and/or intervention are positively impacted from the screening outcome (Hayes, Nelson, & Jarrett, 1987).

Curriculum-Based Measurement

The third assumption of the conceptual framework is that Curriculum-Based Measures are feasible and easy to use screening tools for mathematics within the general education arena. Curriculum-Based Measurement (CBM) was developed at the University of Minnesota Institute for Research on Learning Disabilities (IRLD) to evaluate the Data-Based Program Modification (DBPM), a program designed for special education teachers to use repeated measurement data to formatively evaluate their instruction (Deno & Mirkin, 1977). Conceptually, these procedures were based in Applied Behavior Analysis, Precision Teaching, and psychometrics. Deno and his colleagues believed that if teachers used the data to assess the effectiveness of their instruction, they would become more effective in helping their students gain proficiency in basic skills (Deno, 1992). These procedures, however, were not new to the classroom. At the time, they were known to many educators as Curriculum-Based Assessment. While teachers were familiar with Curriculum-Based Assessment (CBA), Deno and his colleagues felt that CBA was not technically adequate nor did the assessments provide teachers with the opportunity to collect data repeatedly using a single measure (Deno & Fuchs, 1987).

To further understand Curriculum-Based Measurement, the differences between Curriculum-Based Assessment and Curriculum-Based Measurement must be discussed. Fuchs and Deno (1991) stated the difference between Curriculum-Based Assessment and Curriculum-Based Measures is that CBA uses informal assessments that focus on task analysis and criterion-referenced assessment, and CBM uses specific procedures to measure the basic skills of students.

Additionally, CBM provides teachers with specific guidelines on selecting materials to measure student performance so that data outcomes produce reliable and valid student scores. These guidelines, however, do not contain specific measurement protocols. Teachers may employ these techniques using any type of stimulus material based on the needs of their students. So, it is possible for a teacher to apply these procedures to published curricula or general materials. In addition, a teacher can measure a student's progress on the same task over time using alternate-forms of measurement. In contrast, teachers using CBA would analyze a student's performance on a specific skill set.

As CBM research continued to grow, specific psychometric properties and general criteria for Curriculum-Based Measures evolved to include inter-scorer reliability, test-retest reliability, internal consistency reliability, criterion-related validity, and frequent and repeated administration that was time efficient, and cost effective (Deno, 1992). To meet measurement criteria, Deno and Fuchs (1987) identified conditions such as a "behavioral indicators of basic skills" and parameters to measures these behaviors (Deno, Mirkin, & Chang, 1982). The parameters to measure the behaviors included measurement format, length of testing, and type of stimuli. After the behaviors were identified and the alternate measures were designed, evidence of criterion-related validity was investigated through research studies using the measurements (Deno, 1985). As a result, *General Outcome Measuremen*t, GOM, a specific progress monitoring tool, was developed using sources outside of the curriculum (Fuchs & Deno, 1994). This extended the use of CBMs to predicting performance on criterion measures (Good, Simmons, & Kameenui, 2001), developing norms for CBMs (Shinn, 2002), improving teacher instruction (Fuchs, Fuchs, Hamlett & Stecker, 1991), identifying students at-risk for academic failure (Kaminski & Good, 1996) and providing interventions prior to referral for special

education services. Additionally, no formal training was necessary for administration and scoring (Shinn, 1995).

Research in Curriculum-Based Measures initially included parameters of reading and reading related areas such as spelling and written language. Only recently have researchers investigated *General Outcome Measurement* in mathematics. Foegen, Jiban and Deno (2007) conducted a review of the literature in math Curriculum-Based Measures and discovered that 18% of the 578 articles identified were linked to math Curriculum-Based Measurement. Furthermore, only four studies explored measures in early mathematics. A serious imbalance of research is apparent.

Early Numeracy Curriculum-Based Measurement

The final assumption is that prevention and screening is possible through identification of early numeracy skill deficits that lead to math difficulties using CBM. Specifically, in order to correctly identify kindergarteners who may be having math difficulties, a technically adequate screening measure of early numeracy should be administered to all students. Compared to the development of early literacy Curriculum-Based Measures, the development of early mathematics Curriculum-Based Measures significantly lags behind. Currently there are fewer than 10 published studies on early-numeracy Curriculum-Based Measures. See Table 2.1 for a summary of the current early numeracy curriculum-based measures studies. Of those studies some include pre-kindergarten, kindergarten and first grade students or a combination of the two groups (Baglici, Codding, & Tryon, 2010; Chard, et al., 2005; Clarke & Shinn, 2004; Clarke, et al., 2008; Joyce & Wolking, 1987; Lembke & Foegen, 2007; Martinez, et al., 2005; VanDerHeyden, Broussard, George, Lafleur, & Williams, 2011; VanDerHeyden, et al., 2001). The studies also differ in the number of times a year the instruments were administered. In some studies, the measures were administered once a year (Joyce & Wolking, 1987; VanDerHeyden,

et al., 2001), similar to screening measures. In other studies, the measures were administered two to three times a year which is more consistent with Response to Intervention practices as a means of measuring progress or growth over time (Chard, et al., 2005; Clarke & Shinn, 2004; Clarke et. al., 2008; Martinez, et al., 2005). All studies discussed in this literature review claim to measure some aspect of number sense and all measures are individually administered for one minute. Two exceptions are VanDerHeyden, et al. (2001), who administered group early literacy and mathematics Curriculum-Based Measures to kindergarten students, and VanDerHeyden, et al. (2011), who administered an experimental set of early numeracy CBMs class-wide.

Table 2.1
Summary of current early numeracy studies

Study	N	Grade	Measures	Evidence of Reliability/ Validity
Joyce & Wolking (1987)	n/a	K-1	-Count Dots -Name Printed Numbers -Count Backwards from 10	-No reliability or criteria data reported
VanDerHeyden, Witt, Naquin, & Noel (2001)	107	K	-Circle Number -Write Number -Draw Circles	-Alternate-form -CIBS-R
Clarke & Shinn (2004)	52	1	-Oral Counting -Quantity Discrimination -Missing Number -Number Identification	-Alternate-form -Test-retest -*Number Knowledge Test* -WJ III Applied Problems -Add/Sub CBMs
Chard, Clarke, Baker, Otterstedt, Braun, & Katz (2005)	436 438	K 1	-Quantity Discrimination -Missing Number -Number Identification	-No reliability data reported -*Number Knowledge Test*
Clark, Baker, Smolkowski, & Chard (2008)			Oral Counting Quantity Discrimination Missing Number Number Identification Number Writing	
Martinez, Missall, Graney, Aricak, & Clark (2009)	59	K	-Oral Counting -Quantity Discrimination -Missing Number -Number Identification	-Alternate-form -Test-retest -*Stanford 10 Achievement Test* (SAT-10)
Lembke & Foegen (2009)	K-1	382	-Quantity Discrimination -Missing Number -Number Identification -Quantity Array	-Alternate-form -Test-retest -*Test of Early Mathematics Achievement*, Third Edition -*Stanford Early School Achievement Test* -Teacher Rating
Baglici, Codding, & Tryon (2010)	61	K-1	-Oral Counting -Quantity Discrimination -Missing Number -Number Identification	-Alternate-form -AIMSweb 1st grade computation -Report card grades -Teacher rating
VanDerHeyden, Brossouard, Synder, George, Lafleur, & Williams (2011)	46	K	-Pattern Completion -Shape Completion -Comparison of sets with unequal and equal-sized items -adding and taking away objects -Subitivity -Missing Number -Circle Number -Write Number -Draw Circles	-*Inter-scorer agreement* -*Test-retest* -*Test of Early Mathematics Achievement*, Third Edition -First grade CBM probes

While most researchers have agreed that the conceptual idea of number sense needs to be measured in early numeracy CBMs, evidence is still accruing to what measures best represent this concept. Because there is not a universal definition for number sense, researchers are struggling to consistently define the construct. Okamoto and Case (1996), using the *Number Knowledge Test*, explored children's understandings of number and determined that children at different learning stages develop an understanding of a mental counting line. First, children learn to apply a verbal tag to a counted object followed by an ordinal tag to the counted object. Next, they developed the understanding of the cardinal value of a number, and finally, they are able to synthesize the previously learned information and apply it. They described this process as a child's development of number sense. In later research, Kalchman, Moss, and Case (2001) described several components of number sense: fluent ability in the estimation of numbers, ability to discern arbitrary answers, fluent flexibility of mental computation tasks, and fluency and flexibility to recognize and use different representations of numbers. For the purpose of this research study, number sense can be represented using the definition that Kalchman, Moss, and Case (2001) proposed.

Clarke and Shinn (2004) were the first researchers to use the *Number Knowledge Test* (Okamoto & Case, 1996) as a criterion variable. Clarke and Shinn attempted to measure first grade students' number sense using four measures: oral counting, quantity discrimination, missing number and number identification. These measures were referred to as the Test of Early Numeracy (TEN). Clarke, et al. (2008) and Martinez, et al. (2009) administered the same measures to kindergarten students, and Chard, et al. (2005) administered the measures to kindergarten and first grade students. Lembke and Foegen (2009) used a modified version of Clarke and Shinn's (2004) screening instruments. Recently, VanDerHeyden, et al. (2011)

proposed examining number sense class-wide using an experimental set of early numeracy Curriculum-Based Measures: pattern and shape completion, comparison of sets with equal and unequal-sized items, adding and taking away objects, and subitivity.

When investigating the consistency of the early numeracy Curriculum-Based Measures, most researchers report alternate-form reliability coefficients and test-retest reliability coefficients. Clarke and Shinn (2004) results revealed alternate-form reliability correlation coefficients greater than .89 for Oral Counting, Number Identification, and Quantity Discrimination. Correlational coefficients for Missing Number were .78. Martinez et al. (2009) correlation coefficients results indicated Number Identification was the strongest (.92) with Quantity Discrimination and Missing Number greater than .77. Lembke and Foegen (2009) found similar results with alternate-form coefficients ranging from .80 to .90, with the exception of Missing Number. These results indicate that Oral Counting, Number Identification, and Quantity Discrimination exhibit evidence to support using the subscales as screening tools. Furthermore, Clarke and Shinn (2004) found test-retest reliability for 13 weeks ranged from .79-.85 for all measures. Martinez, et al. (2009) test-retest results ranged from .80 (Quantity Discrimination) to .92 (Number Identification). Lembke and Foegen (2009) results also indicated that Number Identification had the strongest test-retest correlations using the mean scores of two probes. Previous research does not report if one form, the mean of two forms, or the median of three forms were used to measure test-retest reliability.

Clarke and Shinn (2004) found that Quantity Discrimination had the strongest relationship with the following individual and group criterion measures: the Applications subtest of the *Woodcock-Johnson III*, the *Number Knowledge Test*, and the group criterion computation math-Curriculum-Based Measure. Martinez et al. (2009) found similar results. Quantity

Discrimination was the only variable to predict *Stanford 10 Achievement Test* (SAT-10) scores. Clarke et. al. (2008) also identified a similar relationship. Correlation coefficients indicated a moderate relationship between Quantity Discrimination and Missing Number and the *Stanford Early School Achievement Tes*t (SESAT), a group achievement test. Lembke and Foegen (2009) found that correlation coefficients were stronger in first grade than in kindergarten for all measures. Missing Number demonstrated the strongest relationship with the criterion variable (SESAT) in the fall. While Chard et al. (2005) found overall weak correlation coefficients with the *Number Knowledge Test*, their study provided the framework for future research using early numeracy measures to assess the construct Number Sense.

Predictive validity results revealed that Quantity Discrimination demonstrated the strongest relationship with criterion variables (Clarke and Shinn, 2004). Baglici et al. (2010) found similar results that Quantity Discrimination scores appeared to have the strongest evidence of a single indicator of early numeracy. Lembke and Foegen’s (2009) study results indicated the strongest relationships existed between teacher’s ratings and the TENS subscale, specifically Missing Number in kindergarten. Marinez et al. (2009) examined the technical adequacy of the combination of the early numeracy measures, and the grouping of Quantity Discrimination, Missing Number, and Number Identification demonstrated the greatest reliability, validity and growth rate coefficients. The combination of Quantity Discrimination and Number Identification, however, showed adequate reliability, validity and growth rate coefficients. Overall, Quantity discrimination appears to be an accurate measure of early numeracy abilities.

Summary

In summary, research in the area of early numeracy Curriculum-Based Measures is in the early stages. Current studies show promise of technically adequate measures to identify students

with weaknesses in early numeracy. Researchers, however, are struggling to define the construct number sense and to identify subscales that measure number sense. Given the current research and the subscales, oral counting, number identification, quantity discrimination and missing number, the subscale quantity discrimination appears to be an accurate and consistent measure of early numeracy abilities. While quantity discrimination appeared to have the strongest relationship with the criterion variables, the subscale missing number demonstrated an adequate relationship with the criterion variable, but consistency estimates were not as strong as the quantity discrimination measure. Continued research, however, is necessary to explore the student growth using these measures and group and individual administration of early numeracy Curriculum-Based Measures.

CHAPTER 3

Methodology

Validity is considered the essential component in designing and evaluating a test. The test itself is not considered valid; rather, scores from the test and the interpretation of the scores are the key components of test validity (Messick, 1980). According to the Standards for Educational and Psychological Testing (American Education Research Association, American Psychologist Association & National Council on Measurement in Education, 1999), when evaluating the validity of a test, it is necessary to consider the following five standards: construct of the test, reliable test scores, proper administration and scoring, appropriate interpretation and use of test scores, and identifiable relationships between test standards and criterion related variables.

Purpose of the Study

In this measurement validation study, the purpose was to investigate the psychometric properties of the *PAM Early Numeracy Screening*. The following research questions were examined to meet the purpose of this study:

1. What was the reliability of the scores on the *PAM Early Numeracy Screening*?
2. What was the concurrent criterion-related validity of the *PAM Early Numeracy Screening*?
3. What evidence of construct validity exists for the use of the *PAM Early Numeracy Screening*?

This chapter is divided into two essential parts: the pilot study and the current study. The pilot study section includes the methodology, results, and discussion. The current study methodology section includes description of the population, instrumentation, procedures for data collection, study limitations, and a chapter summary.

An organizational aid is contained in Table 3.1 to differentiate between the Curriculum-Based Measures from Lembke and Foegen (2009), the pilot study, and the current study.

Table 3.1
Comparison of Curriculum-Based Measures for the Early Numeracy Indicators (Lembke & Foegen, 2009), PAM Early Numeracy Screening-Pilot, and PAM Early Numeracy Screening

	Early Numeracy Indicators (Lembke & Foegen, 2009)	*PAM Early Numeracy Screening-Pilot*	*PAM Early Numeracy Screening*
Quantity Discrimination			
Items	(*N*=63)	(*N*=63)	(*N*=63)
Administration	individual	group	group
Time	1 minute	2 minutes	2 minutes
Type of response	recognition (see-say)	recognition (see-circle)	recognition (see-circle)
Missing Number			
Items	(*N*=63)	(*N*=63)	(*N*=63)
Administration	individual	group	group
Time	1 minute	2 minutes	2 minutes
Type of response	recall (see-say)	recognition (see-circle)	recognition (see-circle)
Number Identification			
Items	(*N*=84)	(*N*=15)	(*N*=21)
Administration	individual	group	group
Time	1 minute	2 minutes, 20 seconds	1 minutes, 45 seconds
Type of response	recall (see-say)	recognition (see-circle)	recognition (see-circle)

Pilot Study Methodology and Results

A pilot study which served as the basis for the development of the current study was conducted in Spring 2009. The design, procedures, and results of the pilot study are discussed to set the stage for the methodology of the current study. The purpose of the pilot study was to determine if a relationship existed between the scores of the *Early Numeracy Indicators*

(individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot* (group measure). The following research questions were posed to meet the purpose:

1. What was the reliability of the scores for the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and for the *PAM Early Numeracy Screening-Pilot* (group measure)?
2. What was the inter-rater reliability for each subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009)?
3. What was the relationship between the individual subscales of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot* (group measure)?
4. What was the concurrent validity of the *PAM Early Numeracy Screening-Pilot* (group measure) using the Calculation and Applied Problems subtests of the *Woodcock-Johnson III Tests of Achievement* (Woodcock, McGrew, & Mather, 2001, 2007)?

Population. The population for this study is kindergarten students in the southeastern region of Louisiana. The accessible sample for this study was kindergarten students in a large school district in the Greater New Orleans Metropolitan Area. The school district consisted of 88 schools. As of February 2009 enrollment at the school district stood at 43,979 students: 48.41% female and 51.59% male. As reflected in the diverse ethnic groups that comprised the school district, the population was multicultural: 49.23% Black, 31.86% White, 12.92% Hispanic, 5.23% Asian, and 0.77% American Indian. In the school district, 64.91% of the total population qualified for free lunch, and 10.05% of the total population qualified for reduced lunch. The students in the district were classified as either fully English proficient (91.57%) or limited

English proficient (8.43%). As of February 2009, enrollment at the school district stood at 3,477 kindergarten students. Approximately 82.91 % of the kindergarten population qualified for free or reduced lunch (Louisiana Department of Education, 2009a).

Sampling. Participants were selected for the study using convenience sampling from two elementary schools within the district: School A and School B. All kindergarten students enrolled in these schools were recruited to participate in the study.

Setting. School A is an elementary school which includes preschool, pre-kindergarten, kindergarten, and first through fifth grades and is located in a school district in the Greater New Orleans Metropolitan Area. As of February 2009 enrollment at School A stood at 649 students: 48.69 % female and 51.31% male. As reflected in the diverse ethnic groups that comprised School A, the population was multicultural: 59.48% Black, 16.49% Hispanic, 14.48% White, 9.24% Asian, and 0.31% American Indian. In School A 79.82% of the total school population qualified for free lunch, and 10.63% of the total school population qualified for reduced lunch. The students in School A were classified as either fully English proficient (83.67%) or limited English proficient (16.33%). As of February 2009 there were 87 kindergarten students enrolled in School A. At the time of this study no students in kindergarten had been identified as needing special education services (Louisiana Department of Education, 2009a).

In School A, there were four kindergarten classrooms. One kindergarten classroom was classified as a Spanish Immersion classroom where students received all instruction in Spanish, with the exception of one hour a day of English language arts. Another kindergarten classroom was classified as an English as a Second Language classroom where students received all instruction in English. The remaining two classrooms had no special classification.

School B is a kindergarten center which includes preschool and kindergarten classrooms and is located in a school district in the Greater New Orleans Metropolitan Area. As of February 2009 enrollment at School B stood at 122 students: 40.98 % female and 59.02% male. As reflected in the diverse ethnic groups that comprised School B, the population was multicultural: 54.10% Black, 40.98% White, 2.46% Hispanic, 1.64% Asian, and 0.82% American Indian. In School B 72.13% of the total school population qualified for free lunch, and 8.20% of the total school population qualified for reduced lunch. The students in School B were classified as either fully English proficient (99.18%) or limited English proficient (0.82%). At the time of this study no students in kindergarten had been identified as needing special education services (Louisiana Department of Education, 2009a).

Sample. The sample consisted of 97 participants, 55% male and 45% female, between the ages of 5 years, 4 months and 7 years, 10 months (M=5 years, 7 months, SD = .50). The sample was composed of the following ethnic groups: 51% Black, 28% White, 15% Hispanic, and 6% Asian. The students in the sample were classified as either fully English proficient (89%) or limited English proficient (11%). In the sample 80% of the students qualified for free lunch, and 8% of the students qualified for reduced lunch. Table 3.2 reports the frequencies and percentages of gender, race, and participants from each school who qualified for free or reduced lunch.

Table 3.2
Frequencies and percentages for demographic variables

	School A		School B	
	Frequency	Percentage	Frequency	Percentage
Gender				
Female	27	48%	17	41.5%
Male	29	52%	24	58.5%
Race				
Asian	6	11%	-	-
Black	24	43%	25	61%
Hispanic	14	25%	1	2%
White	12	21%	15	37%
SES				
Free lunch	48	86%	30	73%
Reduced lunch	5	9%	3	7%
Other	3	5%	8	20%
ESL				
ELL[a]	44	21%	-	-
ELS[b]	12	79%	41	100%

[a]English Language Learners. [b]English Language Speakers.

Instruments and procedures. In this section instruments and procedures used to collect data are described. Data were gathered using Curriculum-Based Measures (CBMs) and achievement tests. Data were gathered using the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the *PAM Early Numeracy Screening-Pilot* (group measure), and tests from the *Woodcock-Johnson III Tests of Achievement.* Procedures included discussions with school principals and faculty with regard to the design of the study as well as data collection training sessions with university personnel.

***Early Numeracy Indicators*.** The *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of four measures with two forms each: Quantity Discrimination, Quantity Array, Missing Number, and Number Identification.

The Quantity Discrimination measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of two forms: Form 1 and Form 2. Test-retest and alternate-form reliability coefficients were computed. Test-retest (mean of two forms) reliability coefficients were .85 or greater for Fall and Winter testing. Alternate-form reliability coefficients for Fall, Winter, and Spring testing were .83 or greater. To explore concurrent validity, correlation coefficients were computed using the mean of the *Early Numeracy Indicators* and three criterion measures: the *Test of Early Mathematics Ability*-3 (r = .45), Woodcock-*McGrew-Werder Mini Battery of Achievement* (r = .38) and *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .59). The strongest relation was between the *Early Numeracy Indicators* and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency*. Predictive validity was examined by correlating *Early Numeracy Indicators* fall scores with spring scores from the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .60) and the *Test of Early Mathematics Ability*-3 (r = .35).

Quantity Discrimination Form 1 consists of four pages, one example page of three items and three pages of 63 items (see Appendix A). Quantity Discrimination Form 2 consists of four pages, one example page of three items and three pages of 63 items (see Appendix B). The design of both probes consists of seven rows of three items per page. Alternate-form reliability coefficients were .83 or greater for Fall, Winter, and Spring testing.

Each Quantity Discrimination measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is individually administered for 1 minute. Example items are reviewed with the student prior to administration of the first item on each probe. The examiner says to the student, "Tell me the number that is bigger" (see Appendix C). Each Quantity Discrimination measure of the *Early Numeracy Indicators* (individual measure; Lembke &

Foegen, 2009) is scored based on correct responses and the mean is calculated (See Appendix D). The scores and mean are recorded on the *Early Numeracy Indicators* Screening Booklet cover (See Appendix E).

The Quantity Array measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of two forms: Form 1 and Form 2. Test-retest and alternate-form reliability coefficients were computed. Test-retest (mean of two forms) reliability coefficients were .72 or greater for Fall and Winter testing. Alternate-form reliability coefficients for Fall, Winter, and Spring testing were .74 or greater. To explore concurrent validity, correlation coefficients were computed using the mean of the *Early Numeracy Indicators* and three criterion measures: the *Test of Early Mathematics Ability*-3 (r = .29), Woodcock-*McGrew-Werder Mini Battery of Achievement* (r = .49) and *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .49). Predictive validity was examined by correlating *Early Numeracy Indicators* fall scores with spring scores from the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .53) and the *Test of Early Mathematics Ability*-3 (r = .35).

Quantity Array Form 1 consists of four pages, one example page of three items and three pages of 72 items. Quantity Array Form 2 consists of four pages, one example page of three items and three pages of 72 items. The design of both probes consists of three rows of eight items per page.

Each Quantity Array measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is individually administered for 1 minute. Example items are reviewed with the student prior to administration of the first item on each probe. The examiner says to the student, "Name the number of dots in each array." Each Quantity Array measure of the *Early*

Numeracy Indicators (individual measure; Lembke & Foegen, 2009) is scored based on correct responses and the mean is calculated. The scores and mean are recorded on the *Early Numeracy Indicators* Screening Booklet cover (See Appendix E).

The Missing Number measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of two forms: Form 1 and Form 2. Test-retest and alternate-form reliability coefficients were computed. Test-retest (mean of two forms) reliability coefficients were .84 or greater for Fall and Winter testing. Alternate-form reliability coefficients for Fall, Winter, and Spring testing were .59 or greater. To explore concurrent validity, correlation coefficients were computed using the mean of the *Early Numeracy Indicators* and three criterion measures: the *Test of Early Mathematics Ability*-3 (r = .48), Woodcock-*McGrew-Werder Mini Battery of Achievement* (r = .57) and *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .64). Predictive validity was examined by correlating *Early Numeracy Indicators* fall scores with spring scores from the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .70) and the *Test of Early Mathematics Ability*-3 (r = .34).

Missing Number Form 1 consists of four pages, one example page of three items and three pages of 63 items (see Appendix F). Missing Number Form 2 consists of four pages, one example page of three items and three pages of 63 items (see Appendix G). The design of both probes consists of seven rows of three items per page.

Each Missing Number measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is individually administered for 1 minute. Example items are reviewed with the student prior to administration of the first item on each probe. The examiner presents the student with a series of four sequential numbers. One number is missing from the series and has

been replace with a blank, and the examiner says to the student, "Tell me the number that goes in the blank" (See Appendix H). Each Missing Number measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is scored based on correct responses and the mean is calculated (See Appendix I). The scores and mean are recorded on the *Early Numeracy Indicators* Screening Booklet cover (See Appendix E).

The Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of two forms: Form 1 and Form 2. Test-retest and alternate-form reliability coefficients were computed. Test-retest (mean of two forms) reliability coefficients were .88 or greater for Fall and Winter testing. Alternate-form reliability coefficients for Fall, Winter, and Spring testing were .87 or greater. To explore concurrent validity, correlation coefficients were computed using the mean of the *Early Numeracy Indicators* and three criterion measures: the *Test of Early Mathematics Ability*-3 (r = .52), Woodcock-*McGrew-Werder Mini Battery of Achievement* (r = .49) and *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .61). Predictive validity was examined by correlating *Early Numeracy Indicators* fall scores with spring scores from the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (r = .64) and the *Test of Early Mathematics Ability*-3 (r = .34).

Number Identification Form 1 consists of four pages, one example page of four items and three pages of 84 items (see Appendix J). Number Identification Form 2 consists of four pages, one example page of four items and three pages of 84 items (see Appendix K). The design of both probes consists of seven rows of four items per page.

Each Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is individually administered for 1 minute. Example items are

reviewed with the student prior to administration of the first item on each probe. The examiner says to the student, “Tell me the number” (See Appendix L). Each Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is scored based on correct responses and the mean is calculated (See Appendix M). The scores and mean are recorded on the *Early Numeracy Indicators* Screening Booklet cover (See Appendix E).

PAM Early Numeracy Screening-Pilot. The researcher designed the *PAM Early Numeracy Screening-Pilot*, a group administered set of Curriculum-Based Measures (CBMs), for kindergarten students. It consists of three subscales with two probes each: Quantity Discrimination, Missing Number, and Number Identification. The *PAM Early Numeracy Screening-Pilot* (group measure) is a modified version of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009).

The Quantity Discrimination subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were adapted from Quantity Discrimination Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) with permission from the author. Quantity Discrimination Probe 1 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of nine pages which included eight example items and 63 items (see Appendix N). Quantity Discrimination Probe 2 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of nine pages which included eight example items and 63 items (see Appendix O). The design of both probes consisted of five rows of two items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate

continuation of the probe. A stop sign was placed at the bottom of the second example page and the last page of the probe to indicate a stopping point. Example and individual probe items were formatted in 36-point Microsoft Sans Serif font. To prevent any occurrence of participant learning, the researcher randomized the order of the items using the randomize function in Microsoft Office Excel.

Each Quantity Discrimination probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is administered for 2 minutes in a group. Example items are reviewed with the students prior to administration of the first item on each probe. Students are instructed to circle the bigger number in each box (See Appendix P). Each Quantity Discrimination probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is scored based on correct responses and the mean is calculated. The scores and mean are recorded on the PAM Early Numeracy Summary Score Sheet-Pilot (See Appendix Q).

The Missing Number subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were copied from Missing Number Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). Missing Number Probe 1 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of 13 pages which included six example items and 63 items (see Appendix R). Missing Number Probe 2 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of 13 pages which included six example items and 63 items (see Appendix S). The design of both probes consisted of six rows of items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate continuation of the probe. A stop sign was placed at the bottom of the second example

page and the last page of the probe to indicate a stopping point. Example and individual probe items were formatted in 36-point Microsoft Sans Serif font. To prevent any occurrence of participant learning, the researcher randomized the order of the items using the randomize function in Microsoft Office Excel.

The example and individual probe items on the Missing Number subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of a stem and response set. The stem consisted of four sequential numbers with one number missing. The numbers in the stem increased either by a single digit, by 5, or by 10. The response set included the keyed response and two distractors. When the numbers in the stem increased by a single digit, the distractors were one number greater than and one number less than the numbers in the sequence. When the numbers in the stem increased by 5, the distractors were the last number in the series plus 10 and the last number in the series plus 1. When the numbers in the stem increased by 10, the distractors were the last number in the series plus 1 and the last number in the series plus 5. When zero was the first number of the stem, there were two response sets. One response set included the keyed response, the last number in the series plus 1, and the number 1. When zero or one was the keyed response, the set included zero or one, the last number in the series plus 1, and the number 2. For both Missing Number probes of the *PAM Early Numeracy Screening-Pilot* (group measure) the location of the keyed response and the distractors within the response set were randomized using the randomize function in Microsoft Office Excel.

Each Missing Number probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is administered for 2 minutes in a group. Example items are reviewed with the students prior to administration of the first item on each probe. Students are instructed to circle the number in each box that comes next in the pattern (See Appendix T). Each Missing Number

probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is scored based on correct responses and the mean is calculated. The scores and mean are recorded on the PAM Early Numeracy Summary Score Sheet-Pilot (See Appendix Q).

The Number Identification subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were designed by the researcher. Number Identification Probe 1 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of five pages which included four example items and 15 items (see Appendix U). Number Identification Probe 2 of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of five pages which included four example items and 15 items (see Appendix V). The design of both probes consisted of six rows of items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate continuation of the probe. A stop sign was placed at the bottom of the second example page and the last page to indicate a stopping point. Example and individual probe items were formatted in 36-point Microsoft Sans Serif font.

The example and individual probe items on the Number Identification subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) consisted of a small picture of a common object and a response set. The response set included four numbers: the keyed response and three distractors. The keyed response and distractors were random numbers from zero to 30. These numbers as well as the position of the keyed response and the distractors within the set were randomized using the randomize function in Microsoft Office Excel.

Each Number Identification probe of the *PAM Early Numeracy Screening-Pilo*t (group measure) was administered for 2 minutes, 20 seconds in a group. Example items are reviewed

with the students prior to administration of the first item on each probe. Students are instructed to put a finger on a specific picture and circle the keyed response in the response set. The stimulus for each item is administered every 8 seconds. The examiner allows 10 seconds for each participant to turn the page and locate the next picture (see Appendices W and X). Each Number Identification probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is scored based on correct responses and the mean is calculated. The scores and mean are recorded on the PAM Early Numeracy Summary Score Sheet-Pilot (See Appendix Q).

In summary, the *PAM Early Numeracy Screening-Pilot* (group measure) was a modified version of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). It consisted of three subscales: Quantity Discrimination, Missing Number, and Number Identification and each contained two probes. The number of example and individual items varied in each measure.

***Woodcock-Johnson III Tests of Achievement* (WJ III ACH).** The *Woodcock-Johnson III Tests of Achievement* (Woodcock, McGrew & Mather, 2001, 2007) is a norm-referenced test. It consists of two forms, Form A and Form B, and is comprised of 22 tests which assess five curricular areas: reading, oral language, mathematics, writing, and academic knowledge. In each curricular area, individual test and cluster scores are obtained. Cluster scores are derived from two or more individual test scores. Interpretation of results may have greater validity because the score includes an evaluation of multiple skills (Mather & Woodcock, 2001).

Evidence of validity for the WJ III ACH is based on the *Standards for Educational and Psychological Testing* and includes four areas: reliability of test scores, test content, association with external criterion variables, and internal construct. Content validity is based on guidelines for achievement testing, core curriculum areas identified in federal legislation, and Cattell-Horn-

Carroll Theory (CHC) of cognitive abilities which distinguishes specific achievement skills from broad achievement skills. In connection with CHC theory, WJ III ACH authors have attempted to develop individual tests that measure a narrow ability which demonstrates strong reliability. This creates a solid measure of broad abilities. In order to provide evidence of developmental patterns, WJ III ACH authors explored and reported the relationship between test scores and cluster scores and individuals' cognitive achievement growth and decline over time. Evidence of internal structure was measured by examining the relationship between test scores and the theoretical concept that the test was designed to measure. Furthermore, the relationships between test scores and cluster scores that measure similar and dissimilar skills were examined; thus a stronger relationship between similar skills was expected. Confirmatory factor analyses were computed to examine these relationships. Overall, the WJ III ACH produced evidence of validity in accordance with the *Standards for Educational and Psychological Testing* (McGrew, Schrank, & Woodcock, 2007).

For the purpose of this study, two math subtests, Math Calculation and Applied Problems, were administered in order to achieve a Brief Math Cluster Score. Because of the age of the participants, Math Fluency subtest was not administered. Reliability coefficients reported for the Brief Math Cluster Score were as follows: 5 years (r = .96), 6 years (r = .96), and 7 years (r = .94; McGrew, Schrank, & Woodcock, 2007). The manual for the *Woodcock-Johnson III Tests of Achievement* reports that reliability of cluster scores is greater than reliabilities of individual tests because the cluster scores are derived from multiple tests which have been deemed reliable.

Math Calculation, a measure of math achievement, requires the subject to write digits and execute problems of addition, subtraction, multiplication, division, geometry, trigonometry,

and calculus. All problems are completed in the Subject Response Booklet. Calculation subtest reliability coefficients were reported using split-half coefficients and corrected for test length using the Spearman-Brown Prophecy Formula. Split-half coefficients were reported for the following ages: 5 years (r = .97), 6 years (r = .96), and 7 years (r = .87; McGrew, Schrank, & Woodcock, 2007).

Applied Problems measures a subject's ability to analyze and solve math problems. The examiner reads the problem to the subject, and the subject is required to complete the necessary operation to solve the problem and provide a verbal response. Initial problems require simple calculations; however, as the subject progresses, so does the complexity of the problems. Reliability coefficients were computed using split-half coefficients and corrected for test length using Spearman-Brown corrected correlation. Applied Problems subtest split-half coefficients were reported for the following ages: 5 years (r = .92), 6 years (r = .88), and 7 years (r = .91; McGrew, Schrank, & Woodcock, 2007).

Procedures. The pilot study procedures were divided into two parts: personnel and material preparation and data collection. Personnel preparation included meetings with participating school faculty and university team members. Material preparation included organization of testing resources and Internal Review Board (IRB) procedures for participant confidentiality. Data collection included administration and scoring of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the *PAM Early Numeracy Screening-Pilot* (group measure), and the *Woodcock-Johnson III Tests of Achievement* (Woodcock, McGrew & Mather, 2001, 2007)

***Personnel preparation*.** The researcher met with the principals participating in the pilot study to discuss the design of the study, parental consent forms, participant incentives, and data

collection. At School A the researcher met with participating kindergarten teachers to discuss the dissemination and collection of the parental consent forms, participant incentives, and data collection (see Appendix Y). At School B the principal did not think a meeting with the kindergarten teachers was necessary. Prior to data collection, the researcher returned to the schools to collect school demographics and parental consent forms.

Material preparation. The researcher and a university team member prepared all materials. For confidentiality and data management purposes participant data collection booklets were organized in the following order: the *PAM Early Numeracy Screening-Pilot* (group measure), individual probe score sheets for the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), and a PAM Summary Score Sheet-Pilot. The booklet covers were color- coded to identify group or individual administration. Colored sheets of paper were used as dividers between each group probe. In addition the researcher placed on each booklet cover page child assent stickers which included lines for evaluator initials and evaluator codes.

The researcher and a university team member coded the class rosters according to school, classroom teacher, and participant. Each participant was assigned a five-digit number, only identifiable with the data code key which was kept in a locked cabinet in the researcher's university office. Until time of testing, all booklets, forms, and test protocols were stored in the researcher's university office in boxes marked K Math Study.

Data training. Data collectors consisted of 10 university team members, all of whom had bachelor's and/or master's degrees in the field of education or in a related field of study. Data collectors were experienced in administering and scoring CBMs. The researcher trained data collectors in the administration and scoring of early numeracy CBMs. Directions and scoring for each probe were demonstrated. During the training session, data collectors were

required to demonstrate the administration and scoring of each probe. The researcher completed a fidelity checklist on every data collector, and given two opportunities, those who scored a .95 or greater were allowed to administer assessments in the study (see Appendix Z). Two individuals did not consistently achieve .95 or greater and were assigned to assist with participant supervision.

In a separate training session with five experienced Educational Diagnosticians from the university team, the researcher reviewed the administration and scoring of the Calculation and Applied Problems tests of WJ III ACH. The researcher completed a fidelity checklist on every Educational Diagnostician, and those who scored a .95 or greater were allowed to administer assessments in the study (see Appendix AA). All Educational Diagnosticians qualified to administer assessments in the study.

The researcher met with data collectors to review the design of the study. The researcher assigned each data collector to a three-member team and designated a team leader. The collection of child assessment and the date, time, and meeting locations for the study were discussed. To maintain IRB standards, each data collector chose a four-digit identification code.

Data collection. At each school, data were collected over a three-day period. The researcher and a university team member randomly assigned each participant to an individual or group category. Data collectors administered the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the *PAM Early Numeracy Screening-Pilot* (group measure), and the Calculation and Applied Problems tests of the WJ III ACH to every participant. Order of test administration varied. For participants assigned to the group category, the *PAM Early Numeracy Screening-Pilot* (group measure) was administered first, followed by the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). For participants assigned

to the individual category, the Early *Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) was administered first, followed by the *PAM Early Numeracy Screening-Pilot* (group measure). Participants classified as English Language Learners were administered the directions for *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) in English, and then the examiner stated the directions in Spanish. Participants classified as English Language Learners were administered the *PAM Early Numeracy Screening-Pilot* (group measure) in English, and then the examiner stated the directions in Spanish. For all participants the WJ III ACH was administered in English.

Group administration. Before administering the *PAM Early Numeracy Screening-Pilot* (group measure), data collectors and the teacher arranged the classroom desks so that they were adequately spaced from each other. Student booklets were disseminated, and child assent was collected from each participant. The team leader administered the *PAM Early Numeracy Screening-Pilot* (group measure) in the following order: Quantity Discrimination Probe 1, Quantity Discrimination Probe 2, Missing Number Probe 1, Missing Number Probe 2, Number Identification Probe 1, and Number Identification Probe 2. The remaining team members walked around the classroom to ensure that participants were following directions. The classroom teacher assisted the data collectors by managing classroom behavior. The classroom teacher did not participate in the data collection process. Each data collector scored the measures. All probes were scored based on the number of correct responses. Data collectors tallied the correct responses, recorded the scores on the PAM Summary Score Sheet-Pilot, and calculated and recorded the mean for each measure of the *PAM Early Numeracy Screening-Pilot* (group measure). Any questions regarding scoring were answered by the team leader. Following the completion of probe scoring, the team leader collected and reviewed each participant's booklet to

ensure that all scores had been correctly entered on the PAM Summary Score Sheet-Pilot. The team leader then returned all materials to the researcher. This procedure was followed in all group administrations.

Individual administration. Prior to probe administration, the team leader and the classroom teacher discussed a plan for the order in which participants would be chosen for individual administration. Individual probes were administered in the classroom, the cafeteria, and the computer room. Child assent was collected from each participant prior to probe administration. The *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) were administered in the following order: Number Identification Form 1, Number Identification Form 2, Quantity Discrimination Form 1, Quantity Discrimination Form 2, Missing Number Form 1, and Missing Number Form 2. All probes were scored based on the number of correct responses. Data collectors tallied the correct responses, recorded the scores on the PAM Summary Score Sheet-Pilot, and calculated and recorded the mean for each measure of the *PAM Early Numeracy Screening-Pilot* (group measure).

Any questions regarding scoring were answered by the team leader. Following the completion of probe scoring, the team leader collected and reviewed each participant's booklet to ensure that all scores had been correctly entered on the PAM Summary Score Sheet-Pilot. The team leader then returned all materials to the researcher. This procedure was followed in all individual administrations.

Standardized testing. During the administration of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), university team members who functioned as participant supervisors guided participants who were not being individually tested to the library or computer room for standardized testing. The participant supervisors then wrote the name and

date of birth of each participant on the test protocol and introduced the participant to the Educational Diagnostician. Child assent was collected from each participant prior to the administration of the WJ III ACH. After the Educational Diagnostician completed the assessment, the protocol was scored according to the scoring rules in the WJ III ACH manual, and the raw score was recorded on the protocol. A participant supervisor walked the participant back to the classroom. This process continued until lunch time. At that time, each Educational Diagnostician returned the scored protocols to the researcher.

Data assimilation. At the end of each data collection day, the researcher created a plan for the following day. The researcher transported all materials, and on each day, data collection followed the same procedure. This process continued for four days. On the last day of data collection, the researcher gathered all completed data and stored it in the researcher's university office.

The researcher and a data collector removed all identifying information from the data and entered each participant's raw score on the Calculation and the Applied Problems tests into the WJ III Normative Update Compuscore and Profiles Program (WJ III NU; Schrank & Woodcock, 2007). A Brief Math Cluster score was computed using the raw score from the Calculation and the Applied Problem tests.

The researcher and a graduate student entered into Statistical Package for the Social Sciences (SPSS) software the raw scores from the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot* (group measure) and standard scores from the Calculation and the Applied Problem subtests of the WJ III ACH.

Results of the pilot study. The results section includes descriptive and inferential statistics. Descriptive statistics are reported for the demographics of the pilot study sample as well as for the *PAM Early Numeracy Screening-Pilot* (group measure). Inferential statistics include correlation coefficients and multiple regression.

Descriptive statistics. Descriptive statistics for each subscale on the *PAM Early Numeracy Screening-Pilot* (group measure) were computed. In Table 3.3 the range, mean, and standard deviation are reported for each school and for the sample.

Table 3.3
Descriptive statistics for the PAM Early Numeracy Screening-Pilot

Measure	Range	Mean	Standard Deviation
Quantity Discrimination Probe 1			
School A	4-58	28.39	12.64
School B	4-54	33.29	11.52
Overall	4-58	30.83	12.42
Quantity Discrimination Probe 2			
School A	5-63	33.93	13.64
School B	5-59	33.61	11.31
Overall	5-63	34.10	12.60
Missing Number Probe 1			
School A	3-28	13.68	7.30
School B	3-26	13.61	5.38
Overall	3-28	13.69	6.47
Missing Number Probe 2			
School A	4-30	14.47	7.18
School B	3-26	13.37	5.75
Overall	3-30	14.00	6.52
Number Identification Probe 1			
School A	8-15	14.52	1.35
School B	9-15	14.10	1.46
Overall	8-15	14.35	1.39
Number Identification Probe 2			
School A	6-15	14.46	1.61
School B	7-15	14.29	1.58
Overall	6-15	14.39	1.58

Note. (School A, *N* = 56). (School B, *N* = 41).

Descriptive statistics for each subscale on the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) were computed. In Table 3.4, the range, mean, and standard deviation are reported for each school and for the sample.

Table 3.4
Mean and standard deviations on Early Numeracy Indicators

Measure	Range	Mean	Standard Deviation
Quantity Discrimination Probe 1			
School A	0-42	16.66	10.93
School B	1-41	15.15	10.66
Overall	0-42	16.07	10.77
Quantity Discrimination Probe 2			
School A	0-39	16.66	10.71
School B	0-41	15.71	10.41
Overall	0-41	16.35	10.50
Missing Number Probe 1			
School A	0-25	10.73	6.65
School B	0-22	9.34	4.91
Overall	0-25	10.08	5.99
Missing Number Probe 2			
School A	0-25	10.41	6.15
School B	0-18	8.83	4.57
Overall	0-25	9.84	5.66
Number Identification Probe 1			
School A	1-62	24.51	16.22
School B	2-51	21.24	14.17
Overall	1-62	23.08	15.40
Number Identification Probe 2			
School A	0-61	22.23	15.68
School B	2-46	20.15	13.57
Overall	0-61	21.30	14.76

Note. (School A, $N = 56$). (School B, $N = 41$).

Reliability Evidence. Pilot Study Research Question 1: What was the reliability of the scores for the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and for the *PAM Early Numeracy Screening-Pilot* (group measure)? Pearson product-moment correlation coefficients were computed and analyzed between the two probes on each measure of

the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot* (group measure). Results for both measures of early numeracy are reported in Table 3.5. Alternate-form reliability estimates for the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) were .90 or greater. Alternate-form reliability estimates for the *PAM Early Numeracy Screening-Pilot* ranged from .57 to .86. According to Salvia and Ysseldyke (1981), when evaluating reliability estimates for educational decision making using a screening instrument, a standard of .80 or greater is appropriate. Based on this criterion, scores from all three subscales of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and Quantity Discrimination and Missing Number probes of the *PAM Early Numeracy Screening-Pilot* (group measure) appear to demonstrate adequate parallel form reliability.

Table 3.5
Alternate-form reliability estimates for the Early Numeracy Indicators and the PAM Early Numeracy Screening-Pilot

Measure	Individual (*Early Numeracy Indicators*)	Group (*PAM Early Numeracy Screening-Pilot*)
Quantity Discrimination	.964	.863
Missing Number	.903	.775
Number Identification	.937	.582

Note. (N = 97).

Pilot Study Research Question 2: What was the inter-rater reliability for each subscale of the *PAM Early Numeracy Screening-Pilot* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009)? In order to investigate the inter-rater reliability of scores, the researcher and a data collector rescored one third of the student booklets. All group

and individual probes were recalculated. Pearson product moment correlation was used to explore the relationship between the scores. All individual reliability estimates were .99. Group administration reliability estimates ranged from .91 (Missing Number mean) to .99. Reliability estimates are reported in Table 3.6.

Table 3.6
Correlation coefficients indicating degree of inter-rater agreement

Measures	% of agreement	
	Individual (*Early Numeracy Indicators*)	Group (*PAM Early Numeracy Screening-Pilot*)
Quantity Discrimination		
Form 1	.998	.979
Form 2	.999	.963
Mean	.999	.981
Missing Number		
Form 1	.998	.945
Form 2	.999	.995
Mean	.999	.914
Number Identification		
Form 1	.994	.991
Form 2	.993	.994
Mean	.998	.995

Note. (*N* = 44). Individual correlation coefficients represent *Early Numeracy Indicators* and group correlation coefficients represent *PAM Early Numeracy Screening-Pilot*.

Criterion-related Validity Evidence. Pilot Study Research Question 3: What was the relationship between the individual subscales of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot* (group measure)? Pearson product moment correlation coefficients were computed using the six subscales of the *PAM Early Numeracy Screening-Pilot* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). The results for all correlations were

statistically significant and were greater than or equal to .390 (p < .01, 2-tailed). With the exception of group Number Identification, the correlation coefficient between each set of subscales (i.e. Quantity Discrimination individual Forms 1 and 2 and Quantity Discrimination group Probes 1 and 2) was stronger than the correlation coefficients between different subscales (Quantity Discrimination individual Forms 1 and 2 and Missing Number group Probes 1 and 2). Test format is one possible explanation for the weaker correlations between the Number Identification group probes. Curriculum-Based Measures are designed so that a ceiling will not be achieved; however, this was not the case for the group Number Identification subscales. Because of measurement design, a natural ceiling was built into the test; as a result, every participant stopped at the same item. Therefore, the measure may not have been sensitive enough to discriminate between those participants who had acquired the skill and those who had not; thus, the range of difficulty may have been too narrow. Overall, the strongest correlation coefficients were between Missing Number Probe 1 of the *PAM Early Numeracy Screening-Pilot* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) Missing Number Form 1 (r = .773) and Form 2 (r = .757). Correlation coefficients are reported in Table 3.7.

Table 3.7
Correlation coefficients between the Early Numeracy Indicators and the PAM Early Numeracy Screening-Pilot

Individual	Group					
	QD 1	QD 2	MN 1	MN 2	NI 1	NI 2
QD 1	.676	.718	.614	.548	.398	.460
QD 2	.695	.742	.592	.512	.433	.413
MN 1	.488	.522	.773	.671	.415	.433
MN 2	.539	.579	.757	.664	.460	.526
NI 1	.584	.587	.596	.518	.430	.408
NI 2	.578	.571	.569	.486	.390	.378

Note. (*N* = 97).Correlation is significant at the 0.01 level (2-tailed). For individual and group measures QD 1= Quantity Discrimination form/probe 1, QD = Quantity Discrimination form/probe 2, MN 1 = Missing Number form/probe 1, and MN 2 = Missing Number form/probe 2, NI 1= Number Identification form/probe, and NI 2 = Number Identification form/probe 2.

Concurrent validity. Pilot Study Research Question 4: What was the concurrent validity of the *PAM Early Numeracy Screening-Pilot* (group measure) using the Calculation and Applied Problems subtests of the *Woodcock-Johnson III Tests of Achievement*? A multiple regression analysis was conducted using as the independent variables the mean scores on the Quantity Discrimination, Missing Number, and Number Identification probes of the *PAM Early Numeracy Screening-Pilot* (group measure) and as the criterion the Brief Math Cluster score (Forms A and B) on the *Woodcock-Johnson III Tests of Achievement.* The means, standard deviations, and intercorrelations can be found in Table 3.8.

The means of the Quantity Discrimination, Missing Number, and Number Identification probes of the *PAM Early Numeracy Screening-Pilot* (group measure) significantly predicted the WJ III Brief Math Cluster score (Form A), $F(3,29) = 10.22$, $p < .05$, with only Number

Identification significantly contributing to the prediction. The Beta weights, presented in Table 3.9, suggest that the variable, Number Identification, contributed most to predicting the WJ III ACH Brief Math Cluster Score Form A, but Quantity Discrimination and Missing Number also contributed. Multicollinearity is one possible explanation for Quantity Discrimination and Missing Number not statistically contributing to the multiple regression. Note that tolerance for each of these variables is < .54 (1-.46), indicating that substantial multicollinearity exists in the model. Thus, regression coefficients may not reflect true contribution of the variables. Squared structure coefficients (R_s^2) help to indicate that the three variables contributed to explaining the variation in the WJ III ACH scores with Number Identification explaining 73% of shared variance and Quantity Discrimination explaining 64% of shared variance. The adjusted *R* squared value was .51. This indicates that 51% of the variance in the WJ III ACH Brief Math Cluster score (Form A) was explained by the model.

The means of the Quantity Discrimination, Missing Number, and Number Identification probes on the *PAM Early Numeracy Screening-Pilot* (group measure) significantly predicted the WJ III Brief Math Cluster score (Form B), $F(3,60) = 12.25$, $p < .05$, with Quantity Discrimination and Missing Number significantly contributing to the prediction. One possible explanation for Number Identification not significantly contributing to the regression model may be poor participant performance on Number Identification. The Beta weights presented in Table 3.10 suggest that Quantity Discrimination contributes most to predicting the WJ III ACH Brief Math Cluster Score Form B, but Missing Number and Number Identification also contributed. The adjusted *R* squared value was .389. This indicates that 39% of the variance in the WJ III ACH Brief Math Cluster score (Form B) was explained by the model. Furthermore, the squared

structural coefficient (R_s^2) shows that Quantity Discrimination explains 82% of the variance and Missing Number explains73% of the variance.

Table 3.8
Correlations of CBM Means and WJ III ACH Brief Math Cluster Score

	Individual		Group	
Measures	WJ Form A	WJ Form B	WJ Form A	WJ Form B
Quantity Discrimination	.680**	.682**	.545**	.564**
Missing Number	.594**	.622**	.547**	.534**
Number Identification	.615**	.612**	.612**	.267**

Note.* Correlations significant at $p < .05$ level (2 tailed). **Correlations significant at $p < .01$ (2 tailed). Form A ($N = 33$). Form B ($N = 64$).

Table 3.9 *Simultaneous multiple regression summary for the PAM Early Numeracy Screening-Pilot Quantity Discrimination, Missing Number and Number Identification probe means predicting WJ III ACH scores (Form A)*

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	.279	.239	.218	.760	.578
Missing Number	.560	.469	.223	.800	.64
Number Identification	4.717*	1.499	.466	.854	.729

Note. ($N = 33$). R^2 =. 51; $F(3,29)$ = 10.22 , *$p < .05$

Table 3.10

Simultaneous multiple regression summary for the PAM Early Numeracy Screening-Pilot, Quantity Discrimination, Missing Number and Number Identification probe means predicting WJ III ACH scores (Form B)

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	.454*	.155	.367	.904	.817
Missing Number	.743*	.302	.304	.856	.733
Number Identification	.973	1.310	.075	.427	.182

Note. ($N = 64$). R^2 =. 389; $F(3,60) = 12.748$, $p < .05$

Pilot study summary. In summary, 97 participants were administered the *PAM Early Numeracy Screening-Pilot* (group measure), the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), and the Calculation and Applied Problems subtests of the WJ III ACH. Descriptive and inferential statistics were used for data analyses. Alternate-form reliability coefficients indicated that scores from the Quantity Discrimination, Missing Number, and Number Identification forms of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and Quantity Discrimination and Missing Number probes on the *PAM Early Numeracy Screening-Pilot* (group measure) were reliable. The relationship between the subscales of the *PAM Early Numeracy Screening-Pilot* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) were explored using Pearson product moment correlation coefficients. Results indicated that correlations ranged from .378 to .773. Stronger correlations were noted between similar subscales, with the strongest correlation existing between Missing Number Form 1 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and Missing Number Probe 1 of the PAM *Early Numeracy Screening-Pilot* (group measure; r =.773). To further investigate the psychometric properties of

the *PAM Early Numeracy Screening-Pilot* (group measure), a multiple regression analysis was conducted using the *PAM Early Numeracy Screening-Pilot* (group measure) subscales as the predictor and the Brief Math Cluster score from the WJ III ACH as the criterion variable. Results indicated that the *PAM Early Numeracy Screening-Pilot* (group measure) significantly predicted the Brief Math Cluster score. For WJ III ACH Form A, the *PAM Early Numeracy Screening-Pilot* (group measure) explained 51% of the variance, and for WJ III ACH Form B, the *PAM Early Numeracy Screening-Pilot* (group measure) explained 39%, thus indicating a large effect size.

Implications. Results of the pilot study informed the development of the current study. Based on the results of the pilot study, several modifications to the *PAM Early Numeracy Screening-Pilot* (group measure) were required. In an attempt to increase the reliability of scores, modifications to the Number Identification probes were necessary. Editing included increasing the number of problems and the complexity of the response set and decreasing the amount of time a participant has to respond to the stimulus. Additional revisions included adding a line to the top of every probe page for the examiner's initials to ensure that every page is reviewed for possible correct responses. To further investigate the validity of the *PAM Early Numeracy Screening-Pilot* (group measure), additional criterion measures were required. The researcher decided that a group criterion measure was necessary as well as an additional criterion measure which would allow more opportunities for participants to respond.

Current study methodology

The current study was a measurement validation study. The purpose of this study was to investigate the psychometric properties of the *PAM Early Numeracy Screening*. The following research questions were examined to meet the purpose of this study:

1. What was the reliability of the scores on the *PAM Early Numeracy Screening*?
2. What was the concurrent criterion-related validity of the *PAM Early Numeracy Screening*?
3. What evidence of construct validity existed for the use of the *PAM Early Numeracy Screening*?

Population. The population for this study was kindergarten students in the southeastern region of Louisiana. The accessible sample for this study was kindergarten students in a school district in the Greater New Orleans Metropolitan Area. The school district consists of 19 schools. As of October 2009 enrollment at the school district stood at 9,706 students: 51.85% female and 48.15% male. As reflected in the diverse ethnic groups that comprised the school district, the population was multicultural: 58.57% White, 36.24% Black, 3.88% Hispanic, 1.06% Asian, and 0.25% American Indian. In the school district, 40.8% of the total population qualified for free lunch, and 9.18% of the total population qualified for reduced lunch. The students in the district were classified as either fully English proficient (98.97%) or limited English proficient (1.03%). As of October 2009 enrollment at the school district stood at 3,477 kindergarten students. Approximately 53.93 % of the kindergarten population qualified for free or reduced lunch (Louisiana Department of Education, 2009b).

Sampling. For the current study, participants were selected using convenience sampling from one elementary school within the district. All kindergarten students enrolled in this school were recruited to participate in the study.

Setting and sample. The school is an elementary school which includes pre-kindergarten, kindergarten, and first through fifth grades and is located in the Greater New Orleans Metropolitan Area. There were six kindergarten classrooms in the school. As of October

2009, the school enrollment stood at 621 students: 44.7 7 % female, and 55.23% male. As reflected in the diverse ethnic groups that comprised the school district, the population was multicultural: 73.79% Black, 24.32% White, and 2.90% Hispanic. In the school 69.57% of the total school population qualified for free lunch, and 11.27% of the total population qualified for reduced lunch. The students were classified as either fully English proficient (98.87%) or limited English proficient (1.13%). As of October 2009 there were 107 kindergarten students enrolled in the school and 80.37% of the kindergarten students qualified for free or reduced lunch (Louisiana Department of Education, 2009b). Participants in the study included six kindergarten classrooms with class sizes ranging from 14 to 17 students. There was a total 107 kindergarten students enrolled in school and all but three students participated in the study constituting 104 students in the sample. Of the 104 participants, seven were excluded from the study due to incomplete data. Thus data were examined for 97 participants ranging in age from 5 years, 1 month to 7 years, 2 months. Note the school from which the sample was used did not closely match the population of the school district.

Instruments and procedures. In this section instruments and procedures used to collect data are described. Data were gathered using Curriculum-Based Measures (CBMs), achievement tests, and a teacher questionnaire. The CBMs were the Number Identification probe of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen 2007, 2009) and the *PAM Early Numeracy Screening* (group measure). The achievement measures were the *KeyMath-3 Diagnostic Assessment* (KeyMath-3 DA; Connolly, 2007a) and the Primer mathematics subtest of the *Metropolitan Achievement Tests, Eighth Edition* (MAT8; Pearson, 2000). The *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (Lembke & Foegen, 2009) was used as the teacher questionnaire. Procedures included discussion with the school principal and

faculty with regard to the design of the study as well as data collection training sessions with university personnel.

Early Numeracy Indicators. The Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) consists of two forms: Form 1 and Form 2. Number Identification Form 1 consists of four pages, one example page of four items and three pages of 84 items (see Appendix J). Number Identification Form 2 consists of four pages, one example page of four items and three pages of 84 items (see Appendix K). The design of both probes consists of seven rows of four items per page.

Each Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is individually administered for 1 minute. Example items are reviewed with the student prior to administration of the first item on each probe. The examiner says to the student, "Tell me the number" (See Appendix L). Each Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) is scored based on correct responses, and the mean is calculated (See Appendix M). The scores and mean are recorded on the PAM Early Numeracy Summary Score Sheet (See Appendix CC).

***PAM Early Numeracy Screening*.** The researcher designed the *PAM Early Numeracy Screening* (group measure) for kindergarten students. It consists of three subscales with two probes each: Quantity Discrimination, Missing Number, and Number Identification. The *PAM Early Numeracy Screening* (group measure) is a modified version of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen 2007, 2009) and a revised version of the *PAM Early Numeracy Screening-Pilot* (group measure).

For the purpose of the current study, the following general revisions were made to the *PAM Early Numeracy Screening-Pilot* (group measure): scoring procedures, probe directions,

page format, and student response time. To ensure optimal scoring of the Quantity Discrimination and Missing Number subscales of the *PAM Early Numeracy Screening* (group measure), the researcher added a small line to the upper right corner of every problem page (see Appendix BB). Scorers were instructed to write the total number of correct responses for each page on the scoring line in the upper right corner, add the total number from each page, calculate the mean, and record the scores on the PAM Early Numeracy Summary Score Sheet (See Appendix CC). Specific revisions made to the Number Identification subscale consist of an increase in the number of items for each probe and a decrease in student response time.

The Quantity Discrimination subscale of the *PAM Early Numeracy Screening* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were copied from Quantity Discrimination Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). Quantity Discrimination Probe 1 of the *PAM Early Numeracy Screening* (group measure) consisted of nine pages which included eight example items and 63 items (see Appendix DD). Quantity Discrimination Probe 2 of the *PAM Early Numeracy Screening* (group measure) consisted of nine pages which included eight example items and 63 items (see Appendix EE). The design of both probes consisted of five rows of two items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate continuation of the probe. A stop sign was placed at the bottom of the second example page and the last page of the probe to indicate a stopping point. A scoring line was place in the upper right corner of each probe page to provide a location for the page total of correct responses. Example and individual probe items were formatted in 36-

point Microsoft Sans Serif font. To prevent any occurrence of participant learning, the researcher randomized the order of the items using the randomize function in Microsoft Office Excel.

Each Quantity Discrimination probe of the *PAM Early Numeracy Screening* (group measure) is administered for 2 minutes in a group. Example items are reviewed with the students prior to administration of the first item on each probe. Students are instructed to circle the number that is greater than in each box (See Appendix FF). Each Quantity Discrimination probe of the *PAM Early Numeracy Screening* (group measure) is scored based on correct responses. Scorers are instructed to write the total number of correct responses for each page on the scoring line in the upper right corner, add the total number from each page, calculate the mean, and record the scores on the PAM Early Numeracy Summary Score Sheet (See Appendix CC).

While the Quantity Discrimination subscale of the *PAM Early Numeracy Screening* (group measure) was not altered following the pilot study, based on teacher feedback, a semantic change was made to the directions. Originally, the Quantity Discrimination directions instructed students to "Circle the bigger number." Each teacher who participated in the pilot study, however, suggested changing the word "bigger" to "greater than" in order to mirror the district's curriculum. After reviewing the Louisiana State Department of Education Grade Level Expectations (GLEs) for kindergarten, and in order to remain consistent with district and state curriculum standards, the researcher changed the Quantity Discrimination directions to "Circle the number that is greater than" (Louisiana Department of Education, 2010c).

The Missing Number subscale of the *PAM Early Numeracy Screening* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were copied from Missing Number Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). Missing Number Probe 1 of the *PAM Early*

Numeracy Screening (group measure) consisted of 13 pages which included six example items and 63 items (see Appendix GG). Missing Number Probe 2 of the *PAM Early Numeracy Screening* (group measure) consisted of 13 pages which included six example items and 63 items (see Appendix HH). The design of both probes consisted of six rows of items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate continuation of the probe. A stop sign was placed at the bottom of the second example page and the last page of the probe to indicate a stopping point. A scoring line was place to the upper right corner of each probe page to provide a location for the page total of correct responses. Example and individual probe items were formatted in 36-point Microsoft Sans Serif font. To prevent any occurrence of participant learning, the researcher randomized the order of the items using the randomize function in Microsoft Office Excel.

The example and individual probe items on the Missing Number subscale of the *PAM Early Numeracy Screening* (group measure) consisted of a stem and response set. The stem consisted of four sequential numbers with one number missing. The numbers in the stem increased either by a single digit, by 5, or by 10. The response set included the keyed response and two distractors. When the numbers in the stem increased by a single digit, the distractors were one number greater than and one number less than the numbers in the sequence. When the numbers in the stem increased by 5, the distractors were the last number in the series plus 10 and the last number in the series plus 1. When the numbers in the stem increased by 10, the distractors were the last number in the series plus 1 and the last number in the series plus 5. When zero was the first number of the stem, there were two response sets. One response set included the keyed response, the last number in the series plus 1, and the number one. When zero

or one was the keyed response, the set included zero or one, the last number in the series plus 1, and the number two. For both Missing Number probes of the *PAM Early Numeracy Screening* (group measure) the position of the keyed response and the distractors within the response set were randomized using the randomize function in Microsoft Office Excel.

Each Missing Number probe of the *PAM Early Numeracy Screening* (group measure) is administered for 2 minutes in a group. Example items are reviewed with the students prior to administration of the first item on each probe. Students are instructed to circle the number in each box that completes the sequence (See Appendix II). Each Missing Number probe of the *PAM Early Numeracy Screening* (group measure) is scored based on correct responses. Scorers are instructed to write the total number of correct responses for each page on the scoring line in the upper right corner, add the total number from each page, calculate the mean, and record the scores on the PAM Early Numeracy Summary Score Sheet (See Appendix CC).

The Number Identification subscale of the *PAM Early Numeracy Screening* (group measure) consisted of two probes: Probe 1 and Probe 2. Example and individual probe items for both probes were designed by the researcher. Number Identification Probe 1 of the *PAM Early Numeracy Screening* (group measure) consisted of six pages which included four example items and 21 items (see Appendix JJ). Number Identification Probe 2 of the *PAM Early Numeracy Screening* (group measure) consisted of six pages which included four example items and 21 items (see Appendix KK). The design of both probes consisted of six rows of items separated by a gray line. A pictorial smiley face was placed above the first example item and above the first item on page one to indicate a starting point. Arrows were placed at the bottom of each page to indicate continuation of the probe. A stop sign was placed at the bottom of the second example

page and on the last page to indicate a stopping point. Example and individual probe items were formatted in 36-point Microsoft Sans Serif font.

The example and individual probe items on the Number Identification subscale of the *PAM Early Numeracy Screening* (group measure) consisted of a small picture of a common object and a response set. The response set included four numbers: the keyed response and three distractors. The keyed response set contained the numbers 0 to 20, and every number was represented in each probe. The order in which the numbers appeared within the probe was randomized using the randomize function in Microsoft Office Excel. If the keyed response was 0-10, the distractors were the keyed response plus 10, the keyed response plus 20, and a random number from the set 0-20. If the keyed response was 11-20, the distractors were the keyed response minus 10, the keyed response plus 10, and a random number from the set 0-20. For example, if the keyed response was the number 7, the distractors were 17, 27, and a random number from 0-20 not already included in the response set. If the keyed response was the number 13, the distractors were three, 23, and a random number from 0-20 not already included in the response set. Within each response set, there was no repetition of numbers. The position of the keyed response and the distractors within the response set were randomized using the randomize function in Microsoft Office Excel.

Each Number Identification probe of the *PAM Early Numeracy Screening* (group measure) was administered for 2 minutes, 15 seconds in a group. Example items are reviewed with the students prior to administration of the first item on each probe. Students are instructed to put a finger on a specific picture and circle the keyed response in the response set. The stimulus for each item is administered every 5 seconds. The examiner allows 10 seconds for each participant to turn the page and locate the next picture (see Appendices LL and MM). Each

Number Identification probe of the *PAM Early Numeracy Screening-Pilot* (group measure) is scored based on correct responses and the mean is calculated. The scores and mean are recorded on the PAM Early Numeracy Summary Score Sheet (See Appendix CC).

Modifications were made to the Number Identification subscale of the *PAM Early Numeracy Screening-Pilot* (group measure). To adhere to Louisiana GLE standards for Number Recognition in kindergarten, the stimuli were altered from 15 problems in which the keyed response set included random numbers from 0-15 to 21 problems in which the keyed response set contained the numbers 0 to 20, and every number was represented in each probe. To decrease the likelihood of a participant guessing the keyed response and to clearly determine whether or not a participant knew the keyed response, the researcher changed the distractors within each response set.

In summary, for the purpose of the current study, the researcher modified the *PAM Early Numeracy Screening-Pilot* (group measure) to create the *PAM Early Numeracy Screening* (group measure). General revisions were made to all probes while specific revisions were made to Number Identification subscale with regard to test length, administration time, and student response time. Probe modifications were based on the Spring 2009 pilot study results.

***Keymath-3 Diagnostic Assessment* (KeyMath-3 DA).** The *KeyMath-3 Diagnostic Assessment* (Connolly, 2007a) is a norm-referenced, individually administered measure of basic mathematical concepts based on the principles expressed in the National Council of Teachers of Mathematics (NCTM) *Principles and Standards for School Mathematics* and the KeyMath-3 Essential Resources (KeyMath-3 ER). The KeyMath-3 DA is designed to measure mathematical abilities with the option of current grade level functioning and to provide student progress

monitoring in individuals from the ages of 4 years, 6 months through 21 years (Connolly, 2007b).

The KeyMath-3 DA has two parallel forms: Form A and Form B. Each form consists of 10 subtests with 372 items. Scores from these subtests form three Composite Areas: Basic Concepts, Operations, and Applications. Basic Concepts measures an individual's mathematical knowledge using the following five subtests: Numeration, Algebra, Geometry, Measurement and Data Analysis, and Probability. Numeration measures an individual's basic understanding of numbers. Algebra measures an individual's early awareness of algebra and algebraic concepts. Geometry measures an individual's ability to categorize multi-dimensional shapes. Measurement measures an individual's ability to apply standard and nonstandard measurement units. Data Analysis and Probability measures an individual's ability to synthesize and analyze data. Operations measures an individual's ability to calculate problems of addition, subtraction, multiplication, and division using the following three subtests: Mental Computation and Estimation, Addition and Subtraction, and Multiplication and Division. Applications measures an individual's ability to solve math word problems using the following two subtests: Foundations of Problem Solving and Applied Problem Solving. Foundations of Problem Solving measures an individual's ability to recognize the necessary steps and procedures for word problems. Applied Problem Solving measures an individual's ability to apply the steps and procedures to solve math word problems. Administration time can range from 30 to 90 minutes, and it is intended to be administered by qualified individuals (Connolly, 2007b). The KeyMath-3 DA can be scored by hand or by using the KeyMath-3 DA ASSIST™ Scoring and Reporting System (Pearson, 2007).

The standardization sample for the KeyMath-3 DA was based on the *Current Population Survey, 2004* from the U.S. Bureau of the Census, 2004. It included 3,630 individuals, 3,105 of whom were used for the grade sample that included kindergarten through twelfth grade individuals who were from 4 years, 6 months through 21 years, 11 months. Data were collected at 272 test sites in 45 states from March 2006 through December 2006. Age and grade norm-descriptive statistics were reported for sex, race/ethnicity, SES/parent's education, and geographic region. Special populations for the age and grade norm sample included Specific Learning Disability, Speech/Language Impairment, Intellectual Disability, Emotional/Behavioral Disturbance, Developmental Delay, and Other Health Impairment (multiple disabilities, hearing impairments, orthopedic impairments, other health impairments, autism, deafness/blindness, and traumatic brain injury). A total of 196 individuals with special needs were included in the age and grade standardized sample (Connolly, 2007b).

KeyMath-3 DA reliability coefficients include internal consistency, alternate-form, and test-retest. For internal-consistency reliability data, the mean split-half reliability coefficients by grade in the Fall for the Composite Areas (Basics Concepts, Operations, and Applications) and Total Test for Forms A and B are reported in two categories: kindergarten through grade 5 and grades 6 through 12. Form A reliability coefficients ranged from .85 to .95 at the kindergarten through grade 5 level and from .89 to .98 for grades 6 through 12. Form B reliability coefficients ranged from .87 to .96 at the kindergarten through grade 5 level and from .89 to .97 for grades 6 through 12. Alternate–form reliability is reported by grade, pre-kindergarten through grade 12, for each Composite Area: Basic Concepts (adjusted r = .94), Operations (adjusted r = .93), Applications (adjusted r = .88), and Total Test (adjusted r = .96). Test-retest reliability is reported by grade, pre-kindergarten through grade 12, for each Composite Area: Basic Concepts

(adjusted *r* = .95), Operations (adjusted *r* = .93), Applications (adjusted *r* = .93) and Total Test (adjusted *r* = .97). Overall, the scores for the Composites are reliable enough for decision making and reporting (Connolly, 2007b).

Internal consistency estimates were measured using the split-half method. Split-half reliabilities were reported for each form, subtest, and Composite Area by grade for Fall and for Spring, and the correlation coefficients were adjusted for length using the Spearman-Brown prophecy formula. Form A split-half reliabilities for specific subtests for kindergarten in the Spring ranged from .53 to .84: Numeration (.63), Algebra (.53), Geometry (.84), Measurement (.83), Data Analysis and Probability (.69), Foundations of Problem Solving (.77), and Applied Problem Solving (.82). For Composite Areas, spilt-half reliabilities ranged from .86 to .94: Basic Concepts (.91), Operations (.86), Applications (.86), and Total Test (.94). Form B split-half reliabilities for specific subtests for kindergarten in the Spring ranged from .64 to .82: Numeration (.64), Algebra (.73), Geometry (.82), Measurement (.82), Data Analysis and Probability (.68), Foundations of Problem Solving (.73), and Applied Problem Solving (.69). For Composite Areas, split-half reliabilities ranged from .77 to .93: Basic Concepts (.86), Operations (.87), Applications (.77), and Total Test (.93; Connolly, 2007b) .

Construct and content validity are reported for the KeyMath-3 DA. To investigate construct validity, intercorrelational studies were conducted to examine the relationships between scores on the KeyMath-3 DA and scores on the Key Math Revised, Normative Update: A Diagnostic Inventory of Essential Mathematic (KeyMath-R/Nu), the Kaufman Test of Educational Achievement, Second Edition (KTEA-II), the Iowa Tests of Basic Skills® (ITBS®instrument), the Total Test score on the Measures of Academic Progress, and the Group Mathematic Assessment and Diagnostic Evaluation (GMADE). Composite, individual test, and

total test correlation coefficients were reported and ranged from .63 to .93. Overall, validity correlation coefficients range from a low of .63 to a high of .93. Extensive tables listing validity coefficients for the intercorrelational studies are included in the technical manual (Connolly, 2007b).

To explore content validity, the authors of the KeyMath-3 DA designed an outline of the essential curriculum components derived from state standards and from recommendations of the National Council of Teachers of Mathematics across grades. Items were also analyzed by the test author, publishing staff, consultants, and reviewers (Connolly, 2007b).

In summary, the KeyMath-3 DA provides a comprehensive assessment of mathematical abilities. Standardization, reliability, and validity are adequate for the instrument's intended purpose. For the purpose of the current study, the participants were only administered the five subtests which comprise the Basic Concepts Area (Numeration, Algebra, Geometry, Measurement, and Data Analysis and Probability) and the two subtests which comprise the Applications Area (Foundations of Problem Solving and Applied Problem Solving). In this study, the researcher used the KeyMath-3 DA as a screening instrument and considered its reliability and validity adequate for this purpose.

***Metropolitan Achievement Tests Eighth Edition* (MAT8)**. The MAT8 (Pearson, 2000) is a series of norm-referenced and criterion-referenced standardized achievement tests designed for group administration in reading, language arts, math, science, and social studies. MAT8 is designed to reflect current standards of professional organizations. It is intended for students in kindergarten through twelfth grade and offers a complete battery and a short form. Administration time ranges from 1 hour, 30 minutes to 4 hours, 35 minutes. The standardization sample for MAT8 was reportedly based on the 1990 and 1995 Census of Population and Housing

and on the National Center on Educational Statistics (1997-1998). The sample consisted of 140,000 students: 80,000 in Fall 1999 and 60,000 in Spring 2000 (Pearson, 2000). For the purpose of this study, only the math subtest at the Primer level was administered to the participants.

Test reliability and validity were not reported in the testing manual or on the Pearson Publishing website; however, the MAT8 was reviewed by Salvia and Ysseldyke (2007). For most subtests, internal-consistency coefficients are greater than .80; however, there are some reliability estimates below .80. Reliability estimates for test-retest were also reported and ranged from .43 to .91. The MAT8 authors attempted to relate test items to professional standards and school curricula. Salvia and Ysseldyke (2007) believe that the evidence for validity is limited and recommended that the user determine the appropriateness of content validity. With regard to standardization and reliability for group screening purposes, Salvia and Ysseldyke (2007) report that the MAT8 is an adequate measure.

For kindergarten, MAT8 has two levels: Pre-primer which is administered in the Fall and Primer which is administered in the Spring. For the purpose of this study, only the math subtest at the Primer level was administered. The Mathematics subtest measures an individual's math problem solving abilities, math reasoning, and conceptual knowledge of mathematics. The Primer math subtest consists of two example items and 30 items. While the examiner reads the directions for every problem, the students follow along in booklets. Group administration time is approximately 20 minutes. MAT8 may be scored by hand or with the publisher's computerized scoring program. The following scores may be converted from the MAT8 raw scores: scaled scores, individual and group percentile ranks, grade equivalents, normal-curve equivalents, content-cluster performance categories, p-values, and performance standards (Pearson, 2000).

Raw scores are converted using the *Metropolitan Achievements Tests, Eight Edition Spring Multilevel NORMS BOOK* (Harcourt Educational Measurement, 2001).

***Early Math Measures Study Teacher Rating of Students' Math Proficiency*.** The *Early Math Measures Study Teacher Rating of Students' Math Proficiency* is a teacher survey developed by Lembke and Foegen (2009). Teachers are asked to consider each student's general math proficiency in comparison to grade level peers. Each student is rated on a Likert scale of 1 to 7 with 1 representing students who are the least proficient and 7 the most proficient (see Appendix NN).

Procedures. The current study procedures were divided into three parts: personnel and material preparation, data training, and data collection. Personnel preparation included meeting with the school faculty to discuss the design of the study. Material preparation included organization of testing resources and IRB procedures for participant confidentiality. Data training included the training of university team members and university graduate tests and measurements students on the administration and scoring of the Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the *PAM Early Numeracy Screening* (group measure), the KeyMath-3 DA (Connolly 2007a), and the (Pearson, 2000). Data collection included administration and scoring of the Number Identification measure of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the *PAM Early Numeracy Screening* (group measure), the KeyMath-3 DA (Connolly 2007a), the *Metropolitan Achievement Tests* Eighth Edition (Pearson, 2000)**,** and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (Lembke & Foegen, 2009).

Personnel preparation. The researcher met with the principal and academic coordinator of the participating school to discuss the *PAM Early Numeracy Screening* (group measure) and the proposed study. At a later date, the researcher met with kindergarten teachers and the academic coordinator to discuss the distribution and collection of parental consent forms (see Appendix OO), participant incentives, data collection, and the dissemination of participants' results. During this meeting the researcher provided each teacher with an envelope containing parent consent forms and participant incentives. Prior to data collection, the researcher returned to the school to collect school demographics, parental consent forms, and class rosters.

Material preparation. The researcher and a university team member prepared all materials. For confidentiality and data management purposes participant data collection booklets were organized in the following manner: the *PAM Early Numeracy Screening (*group measure) followed by the individual probe score sheets for Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). The researcher placed a small round sticker on the back cover of every third *PAM Early Numeracy Screening* (group measure) student booklet to indicate that the participant was to be administered Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). In addition, the researcher placed child assent stickers on each booklet cover page which included lines for evaluator initials and evaluator codes.

A university team member coded the class rosters according to school, classroom teacher, and participant. Each participant was assigned a five-digit number only identifiable with a data code key which was kept in a locked cabinet in the researcher's university office. The researcher created computerized labels for each participant's number, and the label was affixed to the cover page of each *PAM Early Numeracy Screening* (group measure) participant booklet, KeyMath-

3DA protocol, and MAT8 protocol. To identify the participant for the examiner, the university team member wrote each participant's name and date of birth on a piece of paper and stapled it to the last page of the KeyMath-3DA protocol. The researcher also wrote each participant's name on the back cover of the *PAM Early Numeracy Screening* participant booklet (group measure). For each classroom, the university team member wrote the participants' names and the teacher's name on the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (Lembke & Foegen 2009). Until time of testing, all booklets, forms, and test protocols were stored in the researcher's university office in a container marked K math study.

Data training. Data collectors consisted of 10 university team members, all of whom had bachelor's and/or master's degrees in the field of education or in a related field of study, four Educational Diagnosticians, and eight tests and measurements graduate students. All university team members and Educational Diagnosticians had experience administering CBMs and had collected data in the PAM early numeracy pilot study. The researcher trained university team members and Educational Diagnosticians in the administration and scoring of the *PAM Early Numeracy Screening* (group measure), Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), the KeyMath-3 DA (Connolly, 2007a), and the MAT8 (Pearson, 2000). The researcher trained tests and measurements graduate students in the administration and scoring of the *PAM Early Numeracy Screening* (group measure), Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009), and the KeyMath-3 DA (Connolly, 2007).

The researcher conducted several data training sessions on the administration and scoring of early numeracy CBMs. For the university team members, the researcher reviewed procedures

for the administration and scoring of the *PAM Early Numeracy Screening* (group measure) and Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). The researcher then completed a fidelity checklist on every university team member, and those who scored .95 or greater were allowed to administer CBMs in the study (see Appendix Z). All university team members qualified to administer early numeracy CBMs in the study.

In separate training sessions, the researcher trained the eight tests and measurements graduate students in the administration and scoring of the *PAM Early Numeracy Screening* (group measure) and Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009). In another training session, the researcher completed a fidelity checklist on every tests and measurements graduate student, and those who scored a .95 or greater were allowed to administer CBMs in the study (see Appendix Z). All tests and measurements graduate students qualified to administer early numeracy CBMs in the study.

The researcher held separate KeyMath-3DA training sessions for Educational Diagnosticians and tests and measurements graduate students. The researcher followed the same protocol for both training sessions. For the KeyMath-3DA, the researcher reviewed the following processes: individual subtest administration, subtest scoring (e.g., converting raw scores into standard scores and composite scores), and subtest and composite score documentation. Each data collector was required to administer the KeyMath-3DA to another data collector while the researcher conducted a fidelity checklist. The researcher completed a fidelity checklist on each Educational Diagnostician and tests and measurements graduate student and, given two opportunities, those who scored .95 or greater were allowed to administer assessments in the

study (see Appendix AA). All Educational Diagnosticians and tests and measurements graduate students qualified to administer assessments in the study.

The researcher held one MAT8 (Pearson, 2000) training session for university team members. The researcher reviewed general test administration and scoring. The researcher completed a fidelity checklist on each university team member and, given two opportunities, those who scored .95 or greater were allowed to administer the MAT8 in the study. All university team members qualified to administer assessments in the study.

At the completion of all training sessions, the researcher created CBM data collector teams which consisted of a team leader and two or three data collectors. Team leaders were chosen based on experience and performance. Each team leader was either a certified Educational Diagnostician or a Master's Degree teacher with at least five years of experience in the administration and scoring of CBMs. The researcher then appointed team members who were responsible for transporting participants to and from the standardized testing location.

In two separate meetings, the researcher met with university team members and with tests and measurements graduate students to review the design of the study. The researcher assigned each data collector to a two or three-member team and designated a team leader. The collection of child assent and the date, time, and meeting location for the study were discussed. For examiner confidentiality purposes and to maintain IRB standards, the researcher instructed each data collector to choose a four-digit identification code. The researcher then instructed all data collectors to write this code on all student booklets or standardized test protocols they administered or scored.

Data collection. Data were collected over fourteen days. On the first day of data collection, the researcher reviewed the plan for the day with the data collectors. Examiners were

assigned to classrooms and given materials. The data collection process began with each team leader accompanying the researcher to a classroom for a brief introduction between the teacher and the team leader. With the classroom teacher present, the team leader reviewed the procedures for data collection and discussed classroom organization in case not all students in the classroom were participating in the study. Each teacher was provided a treasure chest of appropriate trinkets for participants. The team leader returned to the central meeting area, gathered the remaining team members and appropriate materials, and returned to the assigned classroom.

Group administration. Before administering the *PAM Early Numeracy Screening* (group measure), data collectors and the teacher arranged the classroom desks so that they were adequately spaced from each other. Prior to test administration, the team leader showed all participants the treasure chest and explained the incentive process. Participant booklets were disseminated and child assent was collected from each participant. The team leader administered the *PAM Early Numeracy Screening* (group measure) in the following order: Quantity Discrimination Probe 1, Quantity Discrimination Probe 2, Missing Number Probe 1, Missing Number Probe 2, Number Identification Probe 1, and Number Identification Probe 2. The remaining team members walked around the classroom to ensure that the participants were following directions. The classroom teacher assisted the data collectors by managing classroom behaviors. The classroom teacher did not participate in the data collection process. Each data collector scored the measures. All probes were scored based on the number of correct responses. Data collectors tallied the correct responses, recorded the scores on the PAM Early Numeracy Summary Score Sheet, and calculated and recorded the mean for each measure of the *PAM Early Numeracy Screening* (group measure). Any questions regarding scoring were answered by the team leader. Following the completion of probe scoring, the team leader collected and reviewed

each participant's booklet to ensure that all scores had been correctly entered on the PAM Early Numeracy Summary Score Sheet. The team leader then placed each participant's booklet in the K math study container while the researcher placed a mark next to the participant's name on the class roster. This procedure was followed in all group administrations.

After data collectors administered the *PAM Early Numeracy Screening* (group measure) to all participants, the following tests were administered simultaneously: KeyMath-3DA and Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen , 2009). Educational diagnosticians and tests and measurements graduate students individually administered and scored the KeyMath-3DA. Data collectors who were university team members administered and scored the Number Identification Forms 1 and 2 of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) to participants whose booklets had been coded for administration.

Standardized test administration. University team members who functioned as participant supervisors guided individual participants in classroom 1 (followed by 2, 3, 4, 5, and 6) to the resource room and introduced each participant to the Educational Diagnostician. The Educational Diagnostician collected child assent. The KeyMath-3DA was administered and scored according to the KeyMath-3DA manual to each participant in the following order: Numeration, Algebra, Geometry, Measurement, Data Analysis and Probability, Foundations of Problem Solving, and Applied Problem Solving. Educational Diagnosticians scored all subtests and recorded raw scores. After each participant's score was recorded on the KeyMath-3DA protocol, the participant supervisor walked each participant back to the classroom. This process continued until lunch time and/or until a designated stop time. At that time, each Educational Diagnostician returned the scored protocols to the researcher.

Individual administration. Prior to probe administration, the team leader and the classroom teacher discussed a plan for the order in which participants would be chosen for individual administration. Individual probes were administered in a separate classroom. Child assent was collected from each participant prior to probe administration. The *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) were administered in the following order: Number Identification Form 1 and Number Identification Form 2. All probes were scored based on the number of correct responses. Data collectors tallied the correct responses and recorded the scores on the PAM Early Numeracy Summary Score Sheet. The mean for Number Identification measures of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) was calculated and recorded on the PAM Early Numeracy Summary Score Sheet. Following the completion of probe scoring, the team leader collected and reviewed each participant's booklet to ensure that all scores had been correctly entered on the PAM Early Numeracy Summary Score Sheet. The team leader then placed each participant's booklet in the K math study container while the researcher placed a mark next to the participant's name on the class roster.

Data Collection Process. At the end of data collection on day one, the researcher gave to each teacher a questionnaire entitled *Early Math Measures Study Teacher Rating of Students' Math Proficiency* (Lembke and Foegen, 2009). Teachers were instructed to complete the form according to the directions, and the examiner collected the forms on day two of data collection. The researcher transported all materials daily.

On day two data collectors concurrently administered to all participants by classroom the MAT8. Team leaders and university data collectors then scored the MAT8 and returned the protocols to the researcher who placed the forms in the K math study container.

University team members who functioned as participant supervisors guided individual participants in classroom 2 (followed by 3, 4, 5, and 6) to the resource room and introduced each participant to either the Educational Diagnostician or to the tests and measurements graduate student. Either the Educational Diagnostician or the tests and measurement graduate student collected child assent. The KeyMath-3DA was administered and scored according to the KeyMath-3DA manual to each participant in the following order: Numeration, Algebra, Geometry, Measurement, Data Analysis and Probability, Foundations of Problem Solving, and Applied Problem Solving. Educational Diagnosticians and tests and measurements graduate students scored all subtests and recorded raw scores. After each participant's score was recorded on the KeyMath-3DA protocol, the participant supervisor walked each participant back to the classroom. This process continued until lunch time and/or until a designated stop time. At that time, each Educational Diagnostician and tests and measurements graduate student returned the scored protocols to the researcher. The researcher transported all materials daily.

On days three and four KeyMath-3DA testing continued. University team members who functioned as participant supervisors guided individual participants who had not been previously tested in classroom 2 (followed by 3, 4, 5, and 6) to the resource room and introduced each participant to either the Educational Diagnostician or to the tests and measurements graduate student. Either the Educational Diagnostician or the tests and measurement graduate student collected child assent. The KeyMath-3DA was administered and scored according to the KeyMath-3DA manual to each participant in the following order: Numeration, Algebra, Geometry, Measurement, Data Analysis and Probability, Foundations of Problem Solving, and Applied Problem Solving. Educational Diagnosticians and tests and measurements graduate students scored all subtests and recorded raw scores on the protocol. After each participant's

score was recorded on the KeyMath-3DA protocol, the participant supervisor walked each participant back to the classroom. This process continued until lunch time and/or until a designated stop time. At that time, each Educational Diagnostician and tests and measurements graduate student returned the scored protocols to the researcher.

Thirteen days following the KeyMath-3DA administration, the researcher and five university team members returned to the school to retest participants in classroom 2 using the *PAM Early Numeracy Screening* (group measure). Prior to testing, student booklets were disseminated and child assent was collected from each participant. A previously designated team leader administered the directions, and the remaining university team members walked around the room. Following CBM administration, the team leader and university team members scored the *PAM Early Numeracy Screening* (group measure). All questions regarding scoring were answered by the team leader. Following the completion of probe scoring, the team leader collected and reviewed each participant's booklet to ensure that all scores were entered on the PAM Early Numeracy Summary Score Sheet. The team leader then placed each participant's booklet in the K math study container while the researcher placed a mark next to the participant's name on the class roster.

Data assimilation. Following data collection, the researcher and a university team member removed all identifying information from the student booklets and protocols. Each participant's raw score on Numeration, Algebra, Geometry, Measurement, Data Analysis and Probability, Foundations of Problem Solving, and Applied Problem Solving subtests of the KeyMath-3DA were entered into the into the KeyMath-3 DA ASSIST™ Scoring and Reporting System (Pearson, 2007). After the raw scores were entered for each KeyMath-3DA subtest, a standard score and a composite score were computed. The researcher converted all of the

participants' raw scores on the MAT8 to scaled scores using tables in the *Metropolitan Achievements Tests, Eight Edition Spring Multilevel NORMS BOOK* (Harcourt Educational Measurement, 2001).The researcher and a graduate student entered all participants' scores into SPSS software including individual and group administered early numeracy probes scores, standard scores for Basic Concepts and Applications Areas of the KeyMath-3DA, scaled scores for the MAT8, and raw scores for the *Early Math Measures Study Teacher Rating of Students' Math Proficiency*.

In summary, over a fourteen day period, data were gathered by trained collectors using Curriculum-Based Measures and group and individual math achievement tests. All materials were coded, and the coding key was secured in the researcher's university office. The researcher's role in data collection consisted of personnel training, materials organization, and data entry.

Chapter Summary

In this chapter, the methodology, results, and discussion of the pilot study as well as the methodology of the current study were presented. For both the pilot and current studies, the researcher trained personnel, oversaw the data collection process, and entered all data for analyses.

The pilot study explored the relationship between the *PAM Early Numeracy Screening* (group measure) and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the Brief Math Cluster of the *Woodcock-Johnson III Tests of Achievement*. Results indicated that for the WJ III ACH Brief Math Cluster score Form A, the *PAM Early Numeracy Screening-Pilot* (group measure) explained 39% of the variance, and for Form B, the *PAM Early Numeracy Screening-Pilot* (group measure) explained 46% of the variance. Additionally, scores

from all three subscales of the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and Quantity Discrimination and Missing Number probes of the *PAM Early Numeracy Screening-Pilot* (group measure) appeared to demonstrate adequate parallel form reliability. A strong relationship was determined between the *PAM Early Numeracy Screening-Pilot* Missing Number Probe 1 and the *Early Numeracy Indicators* Missing Number Form 1 (r = .773) and Form 2 (r = .757).

The current study examined the psychometric properties of the *PAM Early Numeracy Screening* using the Number Identification forms from the *Early Numeracy Screening* (Lembke & Foegen 2007, 2009), the *Early Math Measures Study Teacher Rating of Students' Math Proficiency*, the *KeyMath-3 Diagnostic Assessment*, and the *Metropolitan Achievement Tests*, Eighth Edition (MAT8). To investigate this relationship, the researcher conducted multiple-linear regression analyses.

CHAPTER 4

Results

This chapter presents the results of data analyses to address the research questions of this measurement validation study. The purpose of this study was to investigate the psychometric properties of the *PAM Early Numeracy Screening* (group measure). The following research questions were examined to meet the purpose of this study:

1. What was the reliability of the scores on the *PAM Early Numeracy Screening* (group measure)?
2. What was the concurrent criterion-related validity of the *PAM Early Numeracy Screening* (group measure)?
3. What evidence of construct validity existed for the use of the *PAM Early Numeracy Screening* (group measure)?

Descriptive statistics

Descriptive statistics summarize the Curriculum-Based Measures, standardized measures, and teacher questionnaire. Descriptive statistics for each subscale and the means of the *PAM Early Numeracy Screening* (group measure) test and retest, and the Number Identification subscale of the *Early Numeracy Indicators* (individual measure) are presented in Table 4.1. The sample size, range, mean, and standard deviation are reported for each measure.

Table 4.1
Descriptive Statistics for Curriculum-Based Measures

Measure	*N*	Range	Mean (raw score)	SD
PAM Early Numeracy Screening				
Quantity Discrimination				
Probe 1	97	2-58	30.27	11.19
Probe 2	97	5-62	31.10	11.08
Mean	97	3.50-56.50	30.75	10.77
Missing Number				
Probe 1	96	1-26	11.98	5.61
Probe 2	96	0-27	12.75	4.97
Mean	96	1.50-26.50	12.37	4.89
Number Identification				
Probe 1	95	2-21	19.64	2.71
Probe 2	95	2-21	19.89	2.47
Mean	95	2-21	19.81	2.36
Retest *PAM Early Numeracy Screening*				
Quantity Discrimination				
Probe 1	17	20-61	42.88	12.41
Probe 2	17	20-61	43.00	10.46
Mean	17	20-61	42.94	10.76
Missing Number				
Probe 1	17	7-23	16.29	4.27
Probe 2	17	7-24	15.12	3.95
Mean	17	7-22.50	15.71	3.67
Number Identification				
Probe 1	17	14-21	20.41	1.73
Probe 2	17	17-21	19.94	1.34
Mean	17	16-21	20.18	1.36
Early Numeracy Indicators				
Number Identification				
Probe 1	34	3-44	19.50	10.23
Probe 2	34	4-48	18.88	10.39
Mean	34	3.50-46	19.19	10.14

Descriptive statistics for the *Metropolitan Achievement Tests*, Eighth Edition, the Basic Concepts and Applications Areas of the *Key Math-3 Diagnostic Assessment*, and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* are presented in Table 4.2. The sample size, range, mean, and standard deviation are reported for each measure.

Table 4.2
Descriptive Statistics for standardized measures and teacher questionnaire

Measure	*N*	Range	Mean	SD
Standardized measures				
Metropolitan Achievement Tests, Eighth Edition (scaled score)	88	372-552	457.48	33.03
Key Math-3 Diagnostic Assessment (standard score)				
Basic Concepts	93	55-128	88.89	12.72
Applications	93	55-127	86.06	13.88
Teacher Questionnaire (raw score)				
Early Math Measures Study Teacher Rating of Student' Math Proficiency	80	1-7	5	1.84

Reliability evidence

Research Question: What was the reliability of the scores on the *PAM Early Numeracy Screening*? Pearson product-moment correlation coefficients were computed and analyzed between Probe 1 and Probe 2 of each subscale of the *PAM Early Numeracy Screening* (group measure). Results for all subscales of the *PAM Early Numeracy Screening* (group measure) are displayed in Table 4.3. Probes 1 and 2 of each subscale were correlated at 0.70 and above. This indicated that all subscales demonstrated adequate equivalent forms reliability.

Table 4.3
Correlation coefficients of parallel forms of the PAM Early Numeracy Screening

Group Probes			
	QD Probe 2 (*N* = 97)	MN Probe 2 (*N* = 96)	NI Probe 2 (*N* = 95)
QD Probe 1	.835*		
MN Probe 1		.698*	
NI Probe 1			.764*

Note. $^{*}p < .001$.

Pearson product-moment correlation coefficients were computed and analyzed between each subscale on the *PAM Early Numeracy Screening* (group measure) Probe 1, Probe 2 and the mean and its retest. Results for all subscales of the *PAM Early Numeracy Screening* (group measure) are reported in Table 4.4. For the subscale Quantity Discrimination Probe 1 and Probe 2, correlation coefficients were greater than .70 indicating reasonable retest reliability. For the subscale Missing Number, Probe 1 correlation coefficients exhibited reasonable retest reliability. For the subscale Number Identification, Probe 1 demonstrated strong retest reliability given the correlation of .953. For Number Identification and Missing Number, Probe 2 correlation coefficients were low. One factor that may have contributed to the weak correlation was that the small sample (*N* = 15) did not produce enough variability in the scores.

Table 4.4

Test-retest Reliability Estimates of the PAM Early Numeracy Screening

Group Probes		
	N	*r*
Quantity Discrimination		
Probe 1	16	.693*
Probe 2	16	.833*
QD Mean	16	.812*
Missing Number		
Probe 1	16	.658*
Probe 2	16	.188
MN Mean	16	.495
Number Identification		
Probe 1	15	.953*
Probe 2	15	.00
NI Mean	15	.558**

Note. * $p < .001$. ** $p < .05$.

Cronbach's Alpha was computed to measure the internal consistency reliability of sample scores from the *Metropolitan Achievement Tests*, Eighth Edition, and the Basic Concepts and Applications Areas of the *Key Math-3 Diagnostic Assessment*. The alpha for the *Metropolitan Achievement Tests*, Eighth Edition and the *Key Math-3 Diagnostic Assessment* are reported in Table 4.5. The alphas for the *Metropolitan Achievement Tests*, Eighth Edition and the Basic Concepts of the *Key Math-3 Diagnostic Assessment* were .79 or greater, which indicated that the scores from items have reasonable internal consistency reliability. The alpha for the Applications of the *Key Math-3 Diagnostic Assessment* was .66, which indicated minimally adequate reliability. Internal consistency estimate could not be calculated for the *PAM Early Numeracy Screening* because in a timed administration not all items are attempted.

Table 4.5
Internal consistency estimates for achievement tests

Measure	*N*	Alpha
Metropolitan Achievement Tests, Eighth Edition	41	.828
Key Math-3 Diagnostic Assessment		
Basic Concepts	93	.795
Applications	93	.662

Concurrent criterion-related validity evidence

Research Question: What was the concurrent criterion-related validity of the *PAM Early Numeracy Screening* (group measure)? Pearson product-moment correlation coefficients were computed between the *PAM Early Numeracy Screening* (group measure) and the *Metropolitan Achievement Tests,* Eighth Edition. Results for all subscales of the *PAM Early Numeracy Screening* (group measure) and the *Metropolitan Achievement Tests*, Eighth Edition are reported in Table 4.6. Correlation coefficients for the mean of each subscale were stronger than Probe 1 or Probe 2 subscales, with the Quantity Discrimination mean having the strongest relationship.

Table 4.6
Correlation coefficients between the PAM Early Numeracy Screening and Metropolitan Achievement Tests, Eighth Edition

Measure	*N*	*Metropolitan Achievement Tests, Eighth Edition*
Quantity Discrimination		
Probe 1	87	.649*
Probe 2	87	.593*
Mean	87	.650*
Missing Number		
Probe 1	87	.516*
Probe 2	87	.599*
Mean	87	.607*
Number Identification		
Probe 1	86	.424*
Probe 2	86	.370*
Mean	86	.463*

Note. *$p < .001$.

Pearson product-moment correlation coefficients were computed between the *PAM Early Numeracy Screening* (group measure) and the Basic Concepts and Applications Areas of the *Key Math-3 Diagnostic Assessment*. Results for all subscales of the *PAM Early Numeracy Screening* (group measure) and the *Key Math-3 Diagnostic Assessment* are reported in Table 4.7. While all correlation coefficients were statistically significant, the strongest relationships, given the coefficient of .631, were between the Basic Concepts Area and the *PAM Early Numeracy Screening* (group measure) Quantity Discrimination subscale mean; and given the coefficient of .527, between the Applications Area and the *PAM Early Numeracy Screening* (group measure)

Quantity Discrimination subscale mean. Overall, correlation coefficients between the Basic Concepts Area and the *PAM Early Numeracy Screening* (group measure) were stronger than the correlation coefficients between the Applications Area and the *PAM Early Numeracy Screening* (group measure).

Table 4.7
Correlation coefficients between the PAM Early Numeracy Screening and Key Math-3 Diagnostic Assessment

Probe		*Key Math-3 Diagnostic Assessment*	
	N	Basic Concepts	Applications
Quantity Discrimination			
Probe 1	92	.585*	.467*
Probe 2	92	.624*	.542*
Mean	92	.631*	.527*
Missing Number			
Probe 1	90	.449*	.329*
Probe 2	91	.438*	.318*
Mean	90	.485*	.353*
Number Identification			
Probe 1	91	.456*	.351*
Probe 2	90	.440*	.315*
Mean	90	.459*	.358*

Note. *$p < .001$

Pearson product-moment correlation coefficients were computed between the *PAM Early Numeracy Screening* (group measure) and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency*. Results for all subscales of the *PAM Early Numeracy Screening*

(group measure) and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* are reported in Table 4.8. All the correlation coefficients were moderately significant and were greater than or equal to .440. The strongest relationship existed between Quantity Discrimination Probe 1 and the Teacher Questionnaire given the coefficient of .589.

Table 4.8
Correlation coefficients between the PAM Early Numeracy Screening and the Early Math Measures Study Teacher Rating of Students' Math Proficiency

Probe	*N*	Teacher Questionnaire
Quantity Discrimination		
Probe 1	79	.589*
Probe 2	79	.519*
Mean	79	.578*
Missing Number		
Probe 1	78	.500*
Probe 2	78	.440*
Mean	78	.515*
Number Identification		
Probe 1	77	.526*
Probe 2	77	.539*
Mean	77	.541*

Note. * $p < .001$.

Spearman rho correlation coefficients were computed between the individually administered Number Identification subscale of the *Early Numeracy Indicators* (individual measure) and the group administered Number Identification subscale of the *PAM Early*

Numeracy Screening (group measure). Results for the Number Identification subscale of the *PAM Early Numeracy Screening* (group measure) and the *Early Numeracy Indicators* (individual measure) are reported in Table 4.9. The correlation coefficient for the mean of Probes 1 and 2 was higher than the correlation coefficients for each Probe.

Table 4.9
Correlation coefficients between the Number Identification Probes on the PAM Early Numeracy Screening and the Early Numeracy Indicators

Number Identification	*N*	*Early Numeracy Indicators*
Probe 1	33	.675*
Probe 2	34	.597**
Mean	33	.752*

Note. * $p < .001$. ** $p < .05$.

Construct validity evidence

Research Question: What evidence of construct validity existed for the use of the *PAM Early Numeracy Screening*? This question further explores the relationships between the subscale scores of the *PAM Early Numeracy Screening* as group and the criterion measures. Regression analysis was used to answer this question. Table 4.10 presents the independent variables in this analysis.

Table 4.10

Correlation coefficients of PAM Early Numeracy Screening Subscales means

	PAM Early Numeracy Screening subscales means		
	Quantity Discrimination	Missing Number	Number Identification
Quantity Discrimination		.661*	.427*
Missing Number	.661*		.377*
Number Identification	.427*	.377*	

Note. * $p < .01$.

Multiple regression was conducted using the mean scores on the *PAM Early Numeracy Screening* (group measure) Quantity Discrimination, Missing Number, and Number Identification subscales as independent variables and the *Key Math-3 Diagnostic Assessment* Basic Concepts scores as the criterion. This combination of variables significantly predicted the Basic Concepts scores, $F(3,86) = 22.355$, $p < .05$, with the mean scores of the *PAM Early Numeracy Screening* (group measure) Quantity Discrimination and Number Identification significantly contributing to the prediction. The adjusted R squared value was .438. This indicates that 44% of the variance in the *Key Math-3 Diagnostic Assessment* Basic Concepts scores was explained by the model. The beta weights presented in Table 4.11 suggest that Quantity Discrimination mean scores contributed the most to the *Key Math-3 Diagnostic Assessment* Basic Concepts scores. The squared structural coefficients (R_s^2) further explain the independent variables contribution to the model. The squared structural coefficients show that Quantity Discrimination mean scores help explain 87% of R^2 = 44% that the three variables together explain in the model. Missing Number mean scores explain 52% of the 44%, and Number Identification mean scores explain 48% of 44%. Note that even though while Missing Number and Number Identification mean scores explain variation, it is not unique.

Table 4.11

Simultaneous multiple regression summary for the PAM Early Numeracy Screening subscale means predicting Key Math-3 Diagnostic Assessment Basic Concepts scores

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	.562*	.132	.463	.935	.87
Missing Number	.248	.281	.094	.722	.521
Number Identification	1.225*	.475	.232	.69	.48

Note. (*N* = 90). R^2 = .438; *F*(3,86) = 22.355, $^*p < .05$.

Multiple regression was conducted using the mean scores on the *PAM Early Numeracy Screening* (group measure), Quantity Discrimination, Missing Number, and Number Identification subscales as the independent variables and the *Key Math-3 Diagnostic Assessment* Applications scores as the criterion. This combination of variables significantly predicted the Applications scores, *F*(3,86) = 12.378, $p < .05$. The adjusted *R* squared value was .302. This indicates that 30% of the variance in the *Key Math-3 Diagnostic Assessment* Applications scores was explained by this model. The beta weights and squared structure coefficients (R_s^2 = .92) presented in Table 4.12 suggest that the Quantity Discrimination mean scores significantly contributed to predicting the Applications scores, and the Number Identification and Missing Number mean scores also contributed less so to this prediction. The squared structural coefficients (R_s^2) further explain the independent variables contribution to the model. The squared structural coefficients show that Quantity Discrimination mean scores help explain 92% of R^2 = 30% that the three variables together explain in the model. Missing Number mean scores explain 40% of the 30%, and Number Identification mean scores explain 43% of 30%. Note that even though Missing Number and Number Identification mean scores explain variation, it is not

unique. The results indicate that Quantity Discrimination mean scores could be used almost exclusively to predict *Key Math-3 Diagnostic Assessment* Applications scores.

Table 4.12
Simultaneous multiple regression summary for the PAM Early Numeracy Screening subscale means predicting Key Math-3 Diagnostic Assessment Applications scores

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	.629*	.164	.465	.960	.922
Missing Number	-.040	.349	-.014	.636	.404
Number Identification	1.003	.590	.171	.652	.425

Note. (N = 90). R^2 = .302; $F(3,86)$ = 12.378, $^*p < .05$.

Multiple regression was conducted using the mean scores on the *PAM Early Numeracy Screening* (group measure) Quantity Discrimination, Missing Number, and Number Identification subscales as the independent variables and the *Metropolitan Achievement Tests,* Eighth Edition scores as the criterion. This combination of variables significantly predicted the *Metropolitan Achievement Tests*, Eighth Edition scores, $F(3,82) = 31.332$, $p < .05$ with all three variables significantly contributing to this prediction. While the beta weights in Table 4.13 suggest that Quantity Discrimination mean scores contributed the most to predicting the *Metropolitan Achievement Tests*, Eighth Edition scores, the squared structural coefficients show that Missing Number mean scores (R_s^2 =68%) explain almost as much variation as Quantity Discrimination mean scores (R_s^2 =76%). The adjusted R squared value was .534. This indicates that 53% of the variance in the *Metropolitan Achievement Tests*, Eighth Edition scores was explained by this model.

Table 4.13
Simultaneous multiple regression summary for the PAM Early Numeracy Screening subscale means predicting Metropolitan Achievement Tests, Eighth Edition scores

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	1.228*	.331	.367	.871	.759
Missing Number	2.049*	.679	.295	.822	.676
Number Identification	6.169*	1.858	.265	.633	.401

Note. (*N* = 86). R^2 = .534; F(3,82) = 31.332, $^*p$ < .05.

Multiple regression analysis was conducted using the mean scores on the *PAM Early Numeracy Screening-Pilot* (group measure) Quantity Discrimination, Missing Number, and Number Identification subscales as the independent variables and the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* scores as the criterion. This combination of variables significantly predicted the *Early Math Measures Study Teacher Rating of Students' Math Proficiency* scores, F(3,73) = 20.79, p < .05. The beta weights in Table 4.14 suggest that the mean scores of Quantity Discrimination and Number Identification mean scores significantly contributed to this prediction. The squared structural coefficients show that Quantity Discrimination mean scores explain 74% of the 46% , Missing Number mean scores explain 57% of the 46% and Number Identification mean scores explain 64% of the 46% that the three variables together explain in the model indicating that the subscales explain some of the same variation. The adjusted *R* squared value was .461. This indicated that 46% of the variance in the *Teacher Questionnaire* scores was explained by this model.

Table 4.14
Simultaneous multiple regression summary for the PAM Early Numeracy Screening subscale means predicting and Teacher Questionnaire scores

Variable	*B*	*SEB*	β	R_s	R_s^2
Quantity Discrimination	.055*	.020	.321	.861	.741
Missing Number	.073	.042	.191	.755	.57
Number Identification	.231*	.070	.323	.797	.635

Note. (*N*= 77). R^2 = .461; $F(3,73)$ = 20.79, $^{*}p < .05$.

Summary

In summary, this chapter presented the results of data analyses to address the research questions of this measurement validation study. Results revealed acceptable equivalent forms reliability for Probe 1 and Probe 2 of the *PAM Early Numeracy Screening* (group measure), with the Quantity Discrimination subscale demonstrating the strongest correlation coefficient. Test-retest reliability was adequate for the Number Identification Probe 1 subscale. Reasonable reliability was found for Quantity Discrimination Probes 1 and 2 subscale and Missing Number Form 1 subscale. Number Identification and Missing Number Probes 2 subscales revealed low test-retest reliability estimates.

Results of concurrent criterion-related validity for all subscales of the *PAM Early Numeracy Screening* (group measure) and the *Metropolitan Achievement Tests*, Eighth Edition, revealed stronger correlations for the mean than each subscale, with the Quantity Discrimination subscale mean demonstrating the strongest relationship. Concurrent criterion-related validity correlation coefficients for the *PAM Early Numeracy Screening* (group measure) and the *Key Math-3 Diagnostic Assessment* Basic Concepts Area demonstrated a stronger relationship with all three subscale means than with the Applications Area. The *Early Math Measures Study*

Teacher Rating of Student's Math Proficiency and the *PAM Early Numeracy Screening* (group measure) correlation coefficients were significant but moderate in strength, with the strongest relationship existing between Quantity Discrimination Probe 1 subscale and the teacher questionnaire. Correlation between individual and group Number Identification means demonstrated a strong relationship. Internal consistency estimates for the *Key Math -3 Diagnostic Assessment* and the *Metropolitan Achievement Tests*, Eighth Edition indicated reasonable reliability for all three measures.

Construct validity was explored using the means of the subscales of the *PAM Early Numeracy Screening* (group measure) and the Basic Concepts and Applications Areas scores of the *Key Math -3 Diagnostic Assessment*, *Metropolitan Achievement Tests*, Eighth Edition, and the *Early Math Measures Study Teacher Rating of Student's Math Proficiency*. The *PAM Early Numeracy Screening* (group measure) Quantity Discrimination and Number Identification subscales significantly contributed to explaining variation in the Basic Concepts scores. With regard to the Applications scores, Quantity Discrimination significantly contributed to the prediction. For the *Metropolitan Achievement Tests*, Eighth Edition, all three subscales significantly contributed to the prediction. For the *Early Math Measures Study Teacher Rating of Student's Math Proficiency* scores, Quantity Discrimination and Number Identification subscales significantly contributed to the prediction.

CHAPTER 5

Given the existing state of mathematics in the United States (Fleischman et. al. 2010) and the current push for prevention of academic failure (Fuchs et. al., 2005), this study advances the ability of educators to identify early those students with math weaknesses in order to prevent math failure. Based on standardized test scores and individual test items, Mazzocco and Thompson (2005) were able to predict which students in kindergarten were at risk for math learning disabilities in second and third grade. Typically, students are not assessed for learning disabilities until 9 years of age (Shaywitz, 1998). However, Mazzocco and Thompson's (2005) findings support the need for screening of kindergarten students for math weaknesses to ensure early identification and appropriate intervention. One purpose of this study was to develop a measure that was cost effective and psychometrically sound. Chapter 5 provides interpretation, discussion and implications of this study of the psychometric properties of the *PAM Early Numeracy Screening*.

Discussion

In this study construct validity is based on current views and consensus by major psychological and educational measurement organizations. These views are summarized in the following:

> A sound validity argument integrates various strands of evidence into a coherent account of the degree to which existing evidence and theory support the intended interpretation of test scores for specific uses.… Ultimately, the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing system. This includes evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all examinees…. (AERA et al., 1999, p.17).

This study provides beginning evidence that scores from these measures can be used to validly identify students at-risk for mathematics difficulties. A discussion of this evidence follows.

Reliability Evidence

Reliability estimates for equivalent forms of Quantity Discrimination, Missing Number, and the Number Identification on the *PAM Early Numeracy Screening* were strong (Quantity Discrimination, r = .84; Missing Number, r = .70; Number Identification, r = .76) indicating evidence of parallel forms. Other researchers using individually administered measures of early numeracy (Quantity Discrimination, Missing Number, and Number Identification subscales) found similar results. Lembke and Foegen (2009) found that all subscales of the *Early Numeracy Indicators* had alternate-form reliability of .80 or greater. Baglici, Codding, and Tryon (2010) discovered all three measures (Quantity Discrimination, Missing Number, and Number Identification) had reliability estimates greater than .71. Clark and Shinn (2004) also found that Quantity Discrimination and Number Identification demonstrated high alternate-form reliabilities (greater than .89) and Missing Number demonstrated reliabilities greater than .78. Results for the current study indicate promise for the development of group administered alternate forms that perform similarly to individually administered measures.

Test-retest reliability estimates for the subscales on the *PAM Early Numeracy Screening* indicated that in the current study the Number Identification Probe 1 (r = .95) demonstrated good test-retest reliability, Quantity Discrimination Probe 1(r = .69) and Probe 2 (r = .83) and Missing Number Probe 1(r = .66) demonstrated reasonable test-retest reliability, and Number Identification Probe 2 (r = .00) and Missing Number Probe 2 (r = .19) demonstrated poor reliability. A possible explanation for the low test-retest reliability on Number Identification Probe 2 and Missing Number Probe 2 is participant fatigue. The addition of small breaks may

have increased consistency across testing occasions. Additionally, differences in test administrators and ceiling effects on the Number Identification subscale likely contributed to low test-retest reliability. Lembke and Foegen (2009) found all individually administered measures had greater than .80 test-retest reliabilities, and Clark and Shinn (2004) found all individually administered measures had reasonable reliability given the correlation coefficients of .76 or greater. Additional research is needed to explore ways to improve test-retest reliability for all three subscales of the group administered *PAM Early Numeracy Screening*.

Inter-rater agreement was examined and found to be excellent for both the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) and the *PAM Early Numeracy Screening-Pilot*. Given that the correlations were near 1, inter-rater agreement was not examined further for the *PAM Early Numeracy Screening* for the current study. No additional personnel training or test modifications were necessary for the current study.

In general, the *PAM Early Numeracy Screening* subscales produced scores that demonstrated evidence of reliability, particularly alternate-form reliability, for all three subscales and test-retest reliability for at least one probe form for each of the three subscales. Reasons for poor reliability estimates should be explored for improving reliability estimates and correlations with other measures; however, the evidence for this study and the pilot study indicates promise for a group administered measure.

Construct Validity Evidence

Validity evidence was based on test content, response processes, and relations to other variables. Test content evidence of validity was accrued for the *PAM Early Numeracy Screening* (pilot and current study) through careful alignment with prior research and theory related to the development of number sense in young children as well as through expert review of items for

adequate representation of constructs of number sense. Validity substantiation for response processes resulted from (1) alignment with state standards for kindergarten math vocabulary and content difficulty and (2) changes in response time and distractor type and difficulty. Evidence of criterion-related validity exists based on the combined results of the pilot study and the current study. These results, discussed below, are divided into two parts: an examination of the relationships between the standardized measures and the *PAM Early Numeracy Screening* and an examination of the relationships between the non-standardized measures and the *PAM Early Numeracy Screening*.

standardized measures. Correlation coefficients between the *PAM Early Numeracy Screening* and individually administered standardized measures for the current study and pilot study were moderate in magnitude.

For the pilot study, the group administered measure, the *PAM Early Numeracy Screening-Pilot*, explained nearly as much variation in the *Woodcock-Johnson III Tests of Achievement* scores as the individually administered measure, the *Early Numeracy Indicators* (Lembke & Foegen, 2009). Multiple regression results indicated that all three subscale means of the *PAM Early Numeracy Screening-Pilot* explained over half of the variation in Form A of the *Woodcock-Johnson III Tests of Achievement* Math Cluster scores and almost 40%of the variance in Form B of the *Woodcock-Johnson III Tests of Achievement* Math Cluster. The explained variation was smaller for Form B of the *Woodcock-Johnson III Tests of Achievement* possibly because of low reliability on Number Identification subscale or because of standardized test administration issues. Similar results were evident in the current study with 44% and 30% explained variation in the *KeyMath-3 Diagnostic Assessment* Basic Concepts and Applications scores, respectively. With a group administered standardized measure, the *Metropolitan*

Achievement Tests Eighth Edition, similar results were found with 53% of the variation explained by the combination of Quantity Discrimination, Missing Number, and Number Identification. Across all regression models in the current study linking the *PAM Early Numeracy Screening* with a standardized individual and group measures, a single variable, Quantity Discrimination, explained a greater percentage of variation. The other two variables, Missing Number and Number Identification, contributed to explaining substantial variation but were highly correlated with Quantity Discrimination, thus limiting the amount of unique variation explained by either variable.

These results are supported by the research of Clark and Shinn (2004) and Martinez et al. (2009) in which Quantity Discrimination demonstrated a stronger relationship with standardized measures in comparison to Missing Number and Number Identification. In the pilot study, Quantity Discrimination was moderately correlated with the *Woodcock-Johnson III Tests of Achievement*; however, it was not the strongest independent variable in the regression model. Although Clark and Shinn (2004) administered the WJ III, the current study utilized the *KeyMath-3 Diagnostic Assessment* and *Metropolitan Achievement Tests*, Eighth Edition, two measures not known to be used in any other existing studies on early numeracy.

non-standardized measures. Correlation coefficients between the *PAM Early Numeracy Screening* and the *Early Numeracy Indicators* (individual measure; Lembke & Foegen, 2009) ranged from strong (Quantity Discrimination and Missing Number) to moderate (Number Identification). In the current study, a follow-up examination changes in the Number Identification subscale resulted in strong correlations between the *Early Numeracy Indicators* and the *PAM Early Numeracy Screening*. In addition in the current study, correlation coefficients for all three subscales of the *PAM Early Numeracy Screening* and the *Early Math Measures*

Study Teacher Rating of Students' Math Proficiency (Lembke & Foegen, 2009) were equally moderate. Regression results for the teacher questionnaire indicated that all three subscales of the *PAM Early Numeracy Screening* explained almost half of the variation in teacher ratings with Quantity Discrimination again contributing significantly to the model. Lembke and Foegen (2009) found similar results that teacher ratings were moderately correlated with student performance on the *Early Numeracy Indicators*.

Current Study Implications

Results of the pilot study and the current study demonstrate that group measures of early numeracy skills of kindergarten students are comparable to the results of individual and group standardized and non-standardized measures. Currently in the literature, there is no other known group screening measure of early numeracy using the subscales Quantity Discrimination, Missing Number, and Number Identification. These two statements raise important implications. The most important implication is extensive savings of both time and costs. The typical approach for examining the early numeracy skills of kindergarten students is individual testing via standardized (e.g. WJ III ACH, TEMA, KEY Math 3-DA) and/or informal (CBMs, such as the Early Numeracy Indicators, AIMSweb) measures. To administer a standardized test individually, the examiner must be trained, and have experience or academic testing qualifications. Further, an examiner may spend approximately 30-90 minutes administering and scoring a single exam for an individual student (Connolly, 2007b). To test an entire classroom of 20 kindergarten students will take 10-30 hours of testing and scoring time. While informal measures do not require extensive training, time is still a large factor. To test an individual using the *Early Numeracy Indicators* (Lembke & Foegen, 2009), a teacher or trained individual may spend approximately 12 to 15 minutes per student. To test an entire classroom of 20 kindergarten students will take 5

hours of testing and scoring time. Using individually administered measures, an examiner, with or without training, will spend at least 5 hours screening kindergarten students who are at- risk for early numeracy weaknesses.

There are group standardized measures that assess early math achievement (MAT8, Stanford 10). While the *Metropolitan Achievement Tests* Eighth Edition is a group administered achievement test, the examiner must be trained to administer the test. Furthermore, each item on the mathematics subtest of the *Metropolitan Achievement Tests* Eighth Edition (Primer Level) must be read orally to the students. For an entire class of 20 students, a teacher may spend 30 minutes administering the subtest and another 15 minutes scoring each individual student's protocol, for a total of 5.5 hours (Pearson, 2000). This group standardized measure for math achievement is not a time saving solution.

Unlike the individual and group standardized and non-standardized measures, the *PAM Early Numeracy Screening* is a group administered screening measure of early numeracy which can be administered in 15 minutes and scored in an additional 30 minutes. For an entire class of 20 students, this process may be completed in 45 minutes and requires minimal teacher training. For the *PAM Early Numeracy Screening* as with the *Early Numeracy Indicators*, only the directions for each subscale are read to the students. Using the *PAM Early Numeracy Screening*, a teacher or test evaluator may reduce the screening time by 80% compared to other commercially available instruments. This substantial savings in time corresponds to reduction in monetary costs and a reduction in the loss of instructional time due to screening. Given the cost and time savings, the *PAM Early Numeracy Screening* potentially makes prevention more likely to occur, because teachers may be more inclined to use a screener that is easy and quick to administer to a group and score (Wesson, King & Deno, 1984).

Practical Implications

The *PAM Early Numeracy Screening* was administered in a group of at least 14-17 participants with one administrator and one individual assisting the administrator to ensure the participants followed the directions. The administrator scored the students tests within a 30 minute time frame. These results imply that a classroom teacher could administer and score the *PAM Early Numeracy Screening* within a 60 minutes maximum with little to no assistance. No additional classroom assistance is needed to read the directions or score the measure.

Limitations

There were several limitations related to the study. Most limitations were specifically associated with the sample size and selection and the population. In this study, convenience sampling was used. One problem with using convenience sampling is sampling bias which can make it difficult to assure that the results will generalize to other kindergarten students. Another limitation related to the sample was the age range of the participants. Many educators would argue that a 7 year old should not be considered a kindergartener. An additional limitation was that the sample did not reflect the school district population for race. Item analysis was not performed to assess race or gender bias in this study. Another limitation was related to the *PAM Early Numeracy Screening* Number Identification subscale had a ceiling which is not typical of Curriculum-Based Measures; therefore, the specificity of the test may have been altered.

Future Research

Given the results of this study, it may be possible to reduce the time examiners spend administering the *PAM Early Numeracy Screening*. The Quantity Discrimination subscale scores consistently outperformed Missing Number and Number Identification in explaining the variation in scores from standardized and non-standardized measures in this study. Additionally,

the three subscales were moderately to strongly correlated. Further research is needed to investigate the possibility that Quantity Discrimination alone may be effective in identifying kindergarten students who are at-risk for mathematics difficulty.

A source of validity evidence not explored in this study was internal structure. Future research would include the use of techniques such as Rasch Measurement Theory to explore item function (item difficulty and person ability) and construct relevance. This process may result in a very different set of items, possibly fewer items, which may effectively distinguish between students who are at-risk and those who are not. Another result of the information obtained from Rasch Measurement may be the production of evidence for parallel forms and the establishment of item banks. Additionally, using Rasch measurement for item analysis may provide the researcher the opportunity to further develop the measure for children in first grade. Creating measures for first grade enables the measurement of growth in early numeracy. Once a final set of probes is identified, further research needs to explore predictive validity and cut-score development as well as establish national norms that might be used by educators to identify kindergarten students at risk for math difficulties. A longitudinal study should be conducted to explore the predictability of the *PAM Early Numeracy Screening* to identify later math difficulties.

To ensure the use of the *PAM Early Numeracy Screening* with in a classroom setting, future research should investigate teacher administration and scoring of the instrument, the ease of interpretation and usefulness of results, as well as the acceptability of the measure to teachers and others in the academic community.

Conclusion

The current study furthers the existing body of early numeracy research by providing the beginning evidence of a psychometrically sound instrument to screen kindergarten students in a group setting for early numeracy weaknesses. The results were promising; however, additional research is necessary to refine the subscales and to test their adequacy as predictors of math achievement. Early identification leads to prevention of future difficulties in mathematics, this study and the proposed future research contribute to the resolution.

References

AIMSweb. (2011, August 11). Re: AIMSweb pricing. Retrieved from http://www.aimsweb.com/uploads/pdfs/edf_order_site.pdf

American Education Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association

American Education Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Angoff, W. (1988). Validity: An evolving concept. In Wainer, H. & Braun, H. I. (Eds).,*Test Validity* (pp.19-32).New Jersey: Lawrence Erlbaum Associates.

Baglici, S. P., Codding, R., & Tryon, G. (2010). Extending the research on the tests of early numeracy: Longitudinal analyses over two school years. *Assessment for Effective Intervention*, *35*(2), 89-102.

Berch, D. (2005). Making sense of number sense: Implicaitons for children with mathematical disabilities. *Journal of Learning Disabilities*,*38*(4), 333-339.

Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). *Assessment for Effective Intervention, 30*(2), 3-14.

Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An analysis of early numeracy curriculum-based measurement: Examining the role of growth in student outcomes. *Remedial and Special Education*, *29*(1), 46-57

Clarke, B., & Shinn, M. (2004). A Preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, 33(2), 234-248.

Connolly, A. (2007a). *KeyMath-3 Diagnostic Assessment*. Minneapolis, MN: Pearson.

Connolly, A. (2007b). *KeyMath-3 Diagnostic Assessment Manual*. Minneapolis, MN: Pearson.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council in Education.

Deno, S. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3) 219-232.

Deno, S. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure*, *92*(36), 5-11.

Deno, S. L. (2003). Developments in Curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184-192

Deno, S. L., & Fuch, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children, 19*(8), 1-16.

Deno, S. L., & Mirkin, P. (1977). *Data-based program modification: A Manual*. Reston, VA: Council for Exceptional Children.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982) Identifying valid measures of reading. *Exceptional Children*, *18*, 16-26.

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). *Highlights from PISA 2009: Performance of U.S. 15- Year- Old Students in Reading, Mathematics, and Science Literacy in an International Context* (NCES 2011-004). U.S. Department of Education, National Center for Education Statistics. Washington, DC: US. Government Printing Office.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics :A review of the literature. *The Journal of Special Education 41*(2), 121-139.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*(3), 495-513.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, *57*(6), 488-500.

Fuchs, L. S., & Deno, S. L. (1994). Must instructionally useful performance assessment be based in the curriculum? *Exceptional Children, 61*(1), 15-24.

Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Exceptional Children*, *39*(5), 14-20.

Fuchs, L. S., Fuchs, D., Hamlett, C. L. & Steckler, P. M.(1991). Effects of instrumental use of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28, 617-641.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities, 38* 293-304.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessment. *Journal of School Psychology,* 45, 117-135.

Good, R. H. III, Simmons, D. C., & Kameenui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills

for third- grade high stakes outcomes. *Scientific Studies of Reading*, *5*(3), 257-288.

Greeno, J. G. (1991). Number sense as situated knowing in conceptual domain. *Journal for* Research in Mathematics Education, 223 170-218.

Hayes, S. C., Nelson, R. O. & Jarret, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, *42*, 963-974.

Harcourt Educational Measurement. (2001). *Metropolitan Achievements Tests, Eight Edition Spring Multilevel NORMS BOOK*. United States: Harcourt.

Horowitz, Sheldon. (2005). Response to intervention-Tiers without tears. Retrieved from http://www.ncld.org/at-school/general-topics/parentschool-partnership/response-to-intervention-tiers-without-tears

Howden, H. (1989). Teaching Number Sense. *Arithmetic Teacher 36*(6): 6-11.

Joyce, B.G., & Wolking, W. D. (1987). Standard tests and timed curriculum-based assessments: A comparison of two methods for screening high-risk students. *Journal of Psychoeducational Assessment, 5*, 185-193.

Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational number and functions. In S. Carver & D. Klahr (Eds.), *Cognition and instruction*. Mahwah, NJ: Erlbaum

Kaminski, R.A. & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, *25*, 215-227.

Kratochwill, T. R., Albers, C. A., & Shernoff, E. (2004). School-based interventions. *Child and Adolescent Psychiatric Clinics of North America, 13*, 885-903.

Lembke, E., & Foegen, A. (2009). Identifying early numeracy indicators for kindergarten and first-grade students. *Learning Disabilities Research & Practice, 24*(1), 12-30.

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., Kastberg, D., & Jocelyn, L. (2004). *International outcomes of learning in mathematics literacy and problem solving: PISA 2003 results From the U.S. perspective*. (NCES 2005–003). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Louisiana Department of Education (2009a). *Multiple Statistics For Total Reported Public School Students - February 2009*. Retrieved from http://www.doe.state.la.us/offices/infomanagement/student_enrollment_data.html

Louisiana Department of Education (2009b). *Multiple Statistics For Total Reported Public School Students - October 2009*. Retrieved from http://www.doe.state.la.us/offices/infomanagement/student_enrollment_data.html

Louisiana Department of Education (2010c). *Mathematics Grade Level Expectations*: *Kindergarten*. Retrieved from http://www.doe.state.la.us/lde/uploads/3921.pdf

Martinez, R. S., Missall, K. N., Graney, S. B., Ariack, O. T., & Clark, B. (2009). Technical adequacy of early numeracy curriculum-based measurement in kindergarten. *Assessment for Effective Intervention*, *34*(2) 116-125.

Mather, N., & Woodcock, R. W. (2001). Examiner's Manual. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, *20*(3), 142-155.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). Technical Manual. *Woodcock-Johnson III Normative Update*. Rolling Meadows, IL: Riverside Publishing.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50* (9), 741-749.

Methe, S. A., Hojnoski, R., Clarke, B., Owens, B. B., Lilley, P. K., Politylo, B .C., White, K .M., & Marcotte, A.M. (2011). Innovations and future directions for early numeracy curriculum-based measurement: Commentary on the special series. *Assessment for Effective Intervention*, 36, 200-209.

National Center for Education Statistics (2009). *The Nation's Report Card: Mathematics 2009*. (NCES 2010–451). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Center on the Response to Intervention (2011). *Progress monitoring mastery measures tools chart*. Retrieved from http://www.rti4success.org/progressMonitoringTools

Okamoto, Y. & Case, R., (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the Society for Research in Child Development*, *61*, 27-58.

Pearson (2000). *Metropolitan Achievement Tests*, Eighth Edition. United States: Pearson Education.

Pearson, (2007). KeyMath-3 DA ASSIST™ Scoring and Reporting System. Minneapolis, MN: Pearson.

Salvia, J., Ysseldyke, J. E., (1981). *Assessment in Special and Inclusive Education*, (2nd ed.) Boston, MA: Houghton Mifflin.

Salvia, J., Ysseldyke, J. E., & Bolt. (2007). *Assessment: In Special and Inclusive Education*, (10^{th} ed.) Boston, MA: Houghton Mifflin.

Schrank, F. A., & Woodcock, R. W. (2007). WJ II Normative Update Compuscore and Profiles Program (3.0) [Computer Software]*Woodcock-Johnson III*. Rolling Meadows, IL: Riverside Publishing.

Shaywitz, S.E. (1998). Current concepts: Dyslexia. *New England Journal of Medicine*, *338*(5), 307-312.

Shinn, M. (1995). Best practices in curriculum-based measurement and its using a problem-solving model. In J. Grimes & A. Thomas (Eds.). *Best practices in school psychology, 3.* Silver Spring, MD: National association of School Psychologists.

Shinn, M. (2002). Best practices in using curriculum-based measurement in a problem solving model. In J. Grimes & A. Thomas (Eds.). *Best practices in school psychology, 4.* Silver Spring, MD: National association of School Psychologists.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2010). *ESEA Blueprint for Reform*. Retrieved from http://www2.ed.gov/policy/elsec/leg/blueprint/.

VanDerHeyden, A. M., Broussard, C., George, J., Lafleur, S. M., & Williams, C. (2011). Measurement of kindergartners' understanding of early mathematical concepts. School Psychology Review,40 (2), 296-306.

VanDerHeyden, A. M., Witt, J. C., Naquin, G., & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. *School Psychology Review, 20* (3), 363-382.

Walker, H. M., & Severson, H. (1990). Systemic screening for behavioral disorders (SSBD). Dallas, Texas: Cabium.

Wedl, R. (2005). *Response to Intervention: An alternative to traditional eligibility criteria for students with disabilities*. Retrieved from: http://www.educationevolving.org/pdf/Response_to_Intervention.pdf

Wesson, C. L., King, R., & Deno, S .L. (1984). Direct and frequent measurement: If it's so good for us, why don't we use it? *Learning Disability Quarterly*, *(7)*, 45-48.

Wireless Generation. (2011, September 1). Re: mclass Math pricing. Retrieved from http://www.wirelessgeneration.com/

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock-Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside Publishing.

Appendix A
Quantity Discrimination Form 1
Early Numeracy Indicators

Quantity Discrimination, Spring—1

Examples

3	8

10	5

18	9

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 1—student copy

2	5		3	6		8	10
12	7		8	7		6	13
5	6		5	9		15	14
2	4		18	12		17	9
7	9		8	2		14	5
17	0		4	6		5	10
7	16		6	3		1	4

Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 2—student copy

16	6	0	6	9	7
0	6	13	18	3	5
3	7	7	20	4	1
2	5	10	2	1	17
4	5	2	17	10	9
7	17	6	5	11	13
2	5	6	16	9	4

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 3—student copy

8	2		0	10		11	7
4	9		15	17		2	3
1	4		13	5		8	0
3	14		8	6		1	3
7	9		0	17		1	13
8	7		5	2		20	15
5	4		13	8		10	8

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix B
Quantity Discrimination Form 2
Early Numeracy Indicators

Quantity Discrimination, Spring—2

Examples

3	8

10	5

18	9

Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 1—student copy

1	3	6	4	8	5
7	9	7	20	12	16
8	15	2	4	7	5
7	8	3	9	7	10
3	12	5	6	3	13
5	8	14	13	4	5
9	19	16	9	1	5

Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 2—student copy

4	9	11	10	7	5
4	9	6	18	6	0
9	7	0	2	11	2
9	2	1	0	8	16
11	15	5	7	4	3
7	4	15	8	3	2
17	2	3	2	20	7

Research Institute on Progress Monitoring OSEP Award #H324H030003

Quantity Discrimination, page 3—student copy

3	9	6	4	1	9
7	9	15	2	9	7
6	1	2	10	16	9
18	9	2	4	20	11
9	5	11	9	7	5
7	2	19	11	6	2
5	3	1	10	7	12

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix C

Quantity Discrimination Directions

Early Numeracy Indicators

Short Directions for Quantity Discrimination, Individual Administration: (1 minute)

1. Place the student copy in front of the student.

2. Place the student answer sheet/booklet on a clipboard and position so the student cannot see what the examiner records.

3. Say these specific directions to the student:
"Look at the paper in front of you. In each row there are some boxes with numbers in them." (Point to the first box). **"I want you to tell me the number that is bigger."**

4. Correct Response:
"Good. 8 is bigger than 3." (Point to the second set of boxes in the top row.)

Incorrect Response:
"The number that is bigger is 8. You should have said 8 because 8 is bigger than 3." (Point to the second set of boxes in the top row.)

5. Continue with the other example(s). After the examples, turn to the first page of the student copy of the probe.

6. Say to the student:
"When I say begin, I want you to tell me which number is bigger. Start here and go across the page (demonstrate by pointing). **Try each one. If you come to one that you don't know, I'll tell you what to do. Are there any questions? Put your finger on the first one. Ready, begin."**

7. START YOUR STOPWATCH. If the student fails to attempt (does not give the answer to the first problem) after 3 seconds, tell the student to **"Try the next one."**

8. For at least the first 2 to 3 rows of problems, you may need to prompt the student by pointing to the next box and saying **"Tell me which number is bigger."**

9. On the STUDENT ANSWER SHEET, circle the number that the student says. If the student answers incorrectly, circle his/her response and draw an "X" through the entire problem.
-If the student points to a number, encourage the student to "**SAY**" the number. If the student continues to point, write a "P" above the box/number, and score the problem as incorrect.
-If the student skips a problem or row, circle the entire problem or row.

10. The maximum time for each item is 3 seconds. If a student does not provide an answer within 3 seconds, tell the student to
"Try the next one."

11. If the student comes to the end of the page, turn the page to the next page of problems.

12. At the **end of 1 minute**, say **"Stop,"** and **draw a BRACKET** around the last item administered.

13. Repeat these directions and COMPLETE THE EXAMPLE for probe 2.

Scoring Rules

Rule 1: If a student correctly identifies the number score the item as correct.
Rule 2: If the student states any number other than the item number score the item as incorrect and draw an "X" through the entire problem.
Rule 3: If a student hesitates or struggles with a problem for 3 seconds tell the student to "try the next one," score the item as incorrect, and circle the entire problem.
Rule 4: If a student skips a problem or row, score the problem/row as incorrect and circle the entire problem/row.
Rule 5: If a student points to the number, but does not say the number say to the student, "**Tell me which number is bigger**." If the student points to the number (again), score the problem as incorrect and write a **P (Point)** above the number.

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix D
Quantity Discrimination Answer Sheets
Early Numeracy Indicators

Quantity Discrimination I
Answer Sheet Spring 1

Examples

3	8
10	5
18	9

Directions: **Circle the number the student says. Write a "P" ABOVE the box, if the student POINTS to the number. GIVE CREDIT ONLY if the student SAYS the number. If the student SKIPS the problem, circle the ENTIRE PROBLEM.**

Page 1

2	5	3	6	8	10
12	7	8	7	6	13
5	6	5	9	15	14
2	4	18	12	17	9
7	9	8	2	14	5
17	0	4	6	5	10
7	16	6	3	1	4

Page 2

16	6	0	6	9	7
0	6	13	18	3	5
3	7	7	20	4	1
2	5	10	2	1	17
4	5	2	17	10	9
7	17	6	5	11	13
2	5	6	16	9	4

Page 3

8	2	0	10	11	7
4	9	15	17	2	3
1	4	13	5	8	0
3	14	8	6	1	3
7	9	0	17	1	13
8	7	5	2	20	15
5	4	13	8	10	8

QD -I SP -1

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Quantity Discrimination I
Answer Sheet Spring 2

Examples

3	8
10	5
18	9

<u>Directions</u>: **Circle the number the student says. Write a "P" ABOVE the box, if the student POINTS to the number. GIVE CREDIT ONLY if the student SAYS the number. If the student skips the problem, circle the ENTIRE PROBLEM.**

Page 1

1	3	6	4	8	5
7	9	7	20	12	16
8	15	2	4	7	5
7	8	3	9	7	10
3	12	5	6	3	13
5	8	14	13	4	5
9	19	16	9	1	5

Page 2

4	9	11	10	7	5
4	9	6	18	6	0
9	7	0	2	11	2
9	2	1	0	8	16
11	15	5	7	4	3
7	4	15	8	3	2
17	2	3	2	20	7

Page 3

3	9	6	4	1	9
7	9	15	2	9	7
6	1	2	10	16	9
18	9	2	4	20	11
9	5	11	9	7	5
7	2	19	11	6	2
5	3	1	10	7	12

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Appendix E
Early Numeracy Indicators –Screening Booklet

Early Numeracy Indicators-Screening Booklet

Name:	
Teacher:	
School:	
District:	
Academic Year:	
Grade:	

	Fall	**Winter**	**Spring**
Date			
Number Identification			
Form 1			
Form 2			
Mean			
Quantity Discrimination			
Form 1			
Form 2			
Mean			
Missing Number			
Form 1			
Form 2			
Mean			

RIPM OSEP Award #H324H030003

Appendix F
Missing Number Form 1
Early Numeracy Indicators

Missing Number

Example 1

2	3	4	___

0	___	2	3

10	20	30	___

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Missing Number, page 1—Student copy

3 4 5 ___	2 3 ___ 5	4 ___ 6 7
2 ___ 4 5	___ 6 7 8	7 8 ___ 10
3 ___ 5 6	___ 7 8 9	5 10 15 ___
20 30 40 ___	6 ___ 8 9	1 2 3 ___
___ 2 3 4	7 8 9 ___	6 7 8 ___
1 2 3 ___	35 40 45 ___	6 7 8 ___
___ 4 5 6	5 6 7 ___	20 30 40 ___

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Missing Number, page 2—student copy

6 7 8 ___	5 6 7 ___	7 8 ___ 10
___ 3 4 5	20 25 30 ___	15 20 25 ___
1 2 3 ___	1 ___ 3 4	7 ___ 9 10
1 2 ___ 4	40 50 60 ___	___ 7 8 9
4 ___ 6 7	7 8 9 ___	3 4 5 ___
4 5 ___ 7	___ 3 4 5	1 2 3 ___
0 1 2 ___	1 ___ 3 4	___ 2 3 4

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Missing Number, page 3—student copy

1 ___ 3 4	0 1 2 ___	___ 1 2 3
3 4 5 ___	20 30 40 ___	4 ____ 6 7
7 8 9 ___	1 2 3 ___	25 30 35 ___
5 ___ 7 8	___ 4 5 6	5 6 7 ___
1 2 3 ___	___ 7 8 9	10 15 20 ___
3 4 ___ 6	60 70 80 ___	6 7 8 ____
2 3 4 ___	___ 5 6 7	0 ___ 2 3

Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix G
Missing Number Form 2
Early Numeracy Indicators

Missing Number, Spring—2

Missing Number

Example 1

2 3 4 ___	0 ___ 2 3	10 20 30 ___

Missing Number, page 1—Student copy

4 5 6 ___	5 6 7 ___	6 ___ 8 9
5 6 7 ___	6 7 8 ___	___ 8 9 10
1 ___ 3 4	2 3 4 ___	20 25 30 ___
6 7 8 ___	7 ___ 9 10	___ 3 4 5
25 30 35 ___	0 1 2 ___	30 40 50 ___
___ 4 5 6	5 6 7 ___	7 8 9 ___
1 ___ 3 4	___ 1 2 3	70 80 90 ___

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Missing Number, page 2—student copy

2 3 4 ___	1 2 3 ___	___ 3 4 5
7 8 ___ 10	0 1 2 ___	25 30 35 ___
70 80 90 ___	5 ___ 7 8	2 ___ 4 5
7 8 ___ 10	___ 4 5 6	4 ___ 6 7
___ 4 5 6	2 3 4 ___	2 ___ 4 5
4 5 ___ 7	7 8 9 ___	60 70 80 ___
0 1 2 ___	0 ___ 2 3	___ 3 4 5

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Missing Number, page 3—student copy

6 ___ 8 9	4 5 6 ___	___ 6 7 8
0 1 2 ___	6 7 8 ___	___ 6 7 8
3 4 5 ___	5 10 15 ___	3 4 5 ___
10 20 30 ___	7 8 9 ___	___ 5 6 7
1 ___ 3 4	10 15 20 ___	3 4 5 ___
2 3 ___ 5	3 4 5 ___	6 7 8 ___
2 3 4 ___	___ 1 2 3	30 40 50 ___

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix H
Missing Number Form Directions
Early Numeracy Indicators

Directions for Missing Number, Individual Administration: (1 minute)

1. Place the student copy in front of the student.

2. Place the student answer sheet/booklet on a clipboard and position so the student cannot see what the examiner records.

3. Say these specific directions to the student:
"Look at the paper in front of you. Each box has three numbers and a blank." (Point to the first box). **"What number goes in the blank?"**

4. Correct Response:
"Good. The number is 5." (Point to the second box.)

Incorrect Response:
"The number that goes in the blank is 5. You should have said 5 because 5 comes after 4 (2, 3,4)." (Point to the second box.)

5. Continue with the other example(s). After the examples, turn to the first page of the student copy of the probe.

6. Say to the student:
"When I say begin, I want you to tell me what number goes in the blank in each box. Start here and go across the page (demonstrate by pointing). **Try each one. If you come to one that you don't know, I'll tell you what to do. Are there any questions? Put your finger on the first one. Ready, begin."**

7. **START YOUR STOPWATCH** . If the student fails to attempt the first problem, after 3 seconds tell the student to
"Try the next one."

8. For at least the first 2 to 3 rows of problems, you may need to prompt the student by pointing to the next box and saying
"Tell me the number that goes in the blank."

9. On the student answer sheet/booklet, write the number that the student says in the blank.

10. The maximum time for each item is 3 seconds. If a student does not provide an answer within 3 seconds, tell the student to
"Try the next one."

11. If the student comes to the end of the page, turn the page to the next page of problems.

12. At the **end of 1 minute**, say **"Stop,"** and **draw a BRACKET** around the last item administered.

13. Repeat these directions and COMPLETE THE EXAMPLE for probe 2.

Scoring Rules

Rule 1: If a student correctly identifies the number, write his/her response in the blank and score the item as correct.

Rule 2: If the student states any number other than the item number, record his/her response, score the item as incorrect and draw an "X" through the entire problem.

Rule 3: If a student hesitates or struggles with a problem for 3 seconds tell the student to "try the next one," score the item as incorrect and draw an "X: through the entire problem.

Rule 4: If a student skips a problem or row, circle the problem or row and score the problem as incorrect.

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix I
Missing Number Answer Sheets
Early Numeracy Indicators

Missing Number I Answer Sheet Spring 1

Examples

2 3 4 __	0 __ 2 3	10 20 30 __

<u>Directions</u>:
Write the **number** the **student says** in the **blank**. If the student **responds incorrectly**, **draw** an "**X**" through the **entire problem**. If the student **skips** a **problem** or **row**, **circle** the **entire problem** or **row**.

Page 1

3 4 5 __	2 3 __ 5	4 __ 6 7
2 __ 4 5	__ 6 7 8	7 8 __ 10
3 __ 5 6	__ 7 8 9	5 10 15 __
20 30 40 __	6 __ 8 9	1 2 3 __
__ 2 3 4	7 8 9 __	6 7 8 __
1 2 3 __	35 40 45 __	6 7 8 __
__ 4 5 6	5 6 7 __	20 30 40 __

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Page 2 Spring 1

6 7 8 ___	5 6 7 ___	7 8 ___ 10
___ 3 4 5	20 25 30 ___	15 20 25 ___
1 2 3 ___	1 ___ 3 4	7 ___ 9 10
1 2 ___ 4	40 50 60 ___	___ 7 8 9
4 ___ 6 7	7 8 9 ___	3 4 5 ___
4 5 ___ 7	___ 3 4 5	1 2 3 ___
0 1 2 ___	1 ___ 3 4	___ 2 3 4

Page 3

1 ___ 3 4	0 1 2 ___	___ 1 2 3
3 4 5 ___	20 30 40 ___	4 ___ 6 7
7 8 9 ___	1 2 3 ___	25 30 35 ___
5 ___ 7 8	___ 4 5 6	5 6 7 ___
1 2 3 ___	___ 7 8 9	10 15 20 ___
3 4 ___ 6	60 70 80 ___	6 7 8 ___
2 3 4 ___	___ 5 6 7	0 ___ 2 3

Missing Number I Answer Sheet Spring 2

Examples

2	3	4	__	0	__	2	3	10	20	30	__

Directions:
Write the **number** the **student says** in the **blank.** If the student **responds incorrectly, draw** an **"X"** through the **entire problem.** If the student **skips** a **problem** or **row, circle** the **entire problem** or **row.**

Page 1

4	5	6	__	5	6	7	__	6	__	8	9
5	6	7	__	6	7	8	__	__	8	9	10
1	__	3	4	2	3	4	__	20	25	30	__
6	7	8	__	7	__	9	10	__	3	4	5
25	30	35	__	0	1	2	__	30	40	50	__
__	4	5	6	5	6	7	__	7	8	9	__
1	__	3	4	__	1	2	3	70	80	90	__

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Page 2 Spring 2

2 3 4 ___	1 2 3 ___	___ 3 4 5
7 8 ___ 10	0 1 2 ___	25 30 35 ___
70 80 90 ___	5 ___ 7 8	2 ___ 4 5
7 8 ___ 10	___ 4 5 6	4 ___ 6 7
___ 4 5 6	2 3 4 ___	2 ___ 4 5
4 5 ___ 7	7 8 9 ___	60 70 80 ___
0 1 2 ___	0 ___ 2 3	___ 3 4 5

Page 3

6 ___ 8 9	4 5 6 ___	___ 6 7 8
0 1 2 ___	6 7 8 ___	___ 6 7 8
3 4 5 ___	5 10 15 ___	3 4 5 ___
10 20 30 ___	7 8 9 ___	___ 5 6 7
1 ___ 3 4	10 15 20 ___	3 4 5 ___
2 3 ___ 5	3 4 5 ___	6 7 8 ___
2 3 4 ___	___ 1 2 3	30 40 50 ___

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Appendix J
Number Identification Form 1
Early Numeracy Indicators

Number Identification

Example

7	11	8	47

Number Identification, page 1—Student copy

2	7	0	8
13	47	13	40
48	12	20	6
22	4	51	2
68	13	25	18
9	27	12	36
53	5	11	98

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Number Identification, page 2—Student copy

10	7	46	9
24	1	3	30
9	50	17	12
2	95	12	89
5	7	11	33
43	20	4	99
1	9	5	22

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Number Identification, page 3—Student copy

7	0	15	84
14	8	48	2
40	14	39	15
71	2	4	9
29	9	12	14
8	3	48	20
75	43	3	37

Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix K
Number Identification Form 2
Early Numeracy Indicators

Number Identification

Example

7	11	8	47

Research Institute on Progress Monitoring OSEP Award #H324H030003

Number Identification, page 1—Student copy

31	5	17	16
2	38	14	1
16	0	19	18
66	15	9	24
7	1	47	19
6	96	49	3
50	91	49	19

Research Institute on Progress Monitoring OSEP Award #H324H030003

Number Identification, page 2—Student copy

6	2	9	7
29	52	13	41
15	10	41	6
15	43	31	0
12	13	7	99
76	20	2	6
15	26	6	27

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003

Number Identification, page 3—Student copy

32	10	35	2
5	12	7	45
11	8	14	25
72	7	20	19
4	22	38	2
18	6	24	8
98	3	8	100

Research Institute on Progress Monitoring OSEP Award #H324H030003

Appendix L
Number Identification Directions
Early Numeracy Indicators

Short Directions for Number Identification:
(1 minute)

1. Place the student copy in front of the student.

2. Place the student answer sheet/booklet on a clipboard and position so the student cannot see what the examiner records.

3. Say these specific directions to the student:
"Look at the paper in front of you. There are numbers in boxes." (Point to the first box). **"What number is this?"**

4. Correct Response:
"Good. The number is 7." (Point to the second box.)

Incorrect Response:
"This number is 7. What number?"(Point to the second box.)

5. Continue with the other example(s). After the examples, turn to the first page of the student copy of the probe.

6. Say to the student:
"When I say begin, I want you to tell me what number is in each box. Start here and go across the page (demonstrate by pointing). **Try each one. If you come to one that you don't know, I'll tell you what to do. Are there any questions? Put your finger on the first one. Ready, begin."**

7. **START YOUR STOPWATCH** . If the student fails to attempt (does not give the answer to the first problem) after 3 seconds, tell the student to
"Try the next one."

8. For at least the first 2 to 3 rows of problems, you may need to prompt the student by pointing to the next box and saying
"What number is this?"

9. . The maximum time for each item is 3 seconds. If a student does not provide an answer within 3 seconds, tell the student to
"Try the next one."

10. If the student comes to the end of the page, turn the page to the next page of problems.

12. . At the **end of 1 minute**, say **"Stop,"** and **draw a BRACKET** around the last item administered.

13. Turn to probe 2. **"Now you are going to do the same thing. Tell me the number in each box. Do you have any questions? Ready. Begin."** Administer the examples.

Scoring Rules

Rule 1: If a student correctly identifies the number score the item as correct.

Rule 2: If the student states any number other than the item number score the item as incorrect and write the number that the student says above the box.

Rule 3: If a student hesitates or struggles with a problem for 3 seconds tell the student to "try the next one," score the item as incorrect and circle the entire problem.

Rule 4: If a student skips a problem or row, score the problem or row as incorrect, and circle the entire problem or row.

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix M

Number Identification Answer Sheets

Early Numeracy Indicators

Number Identification I

Answer Sheet Spring 1

Examples

7
11
8
47

<u>Directions</u>: **Write incorrect number** responses **ABOVE** the **BOX** and **draw** an **"X" through** the **entire problem**. If the student **skips** a **number**, **circle** the **number.**

Page 1

2	7	0	8
13	47	13	40
48	12	20	6
22	4	51	2
68	13	25	18
9	27	12	36
53	5	11	98

Page 2

10	7	46	9
24	1	3	30
9	50	17	12
2	95	12	89
5	7	11	33
43	20	4	99
1	9	5	22

Page 3

7	0	15	84
14	8	48	2
40	14	39	15
71	2	4	9
29	9	12	14
8	3	48	20
75	43	3	37

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

NI-I SP - 1

Number Identification I
Answer Sheet Spring 2

Examples

7
11
8
47

Directions: Write incorrect number responses **ABOVE the BOX** and **draw** an "X" **through the entire problem**. If the student **skips a number, circle the number.**

Page 1

31	5	17	16
2	38	14	1
16	0	19	18
66	15	9	24
7	1	47	19
6	96	49	3
50	91	49	19

Page 2

6	2	9	7
29	52	13	41
15	10	41	6
15	43	31	0
12	13	7	99
76	20	2	6
15	26	6	27

Page 3

32	10	35	2
5	12	7	45
11	8	14	25
72	7	20	19
4	22	38	2
18	6	24	8
98	3	8	100

OSEP Award #H324H030003
modified by
UNO PAM-RTI/PBS March 2009

Appendix N
Quantity Discrimination Probe 1
PAM Early Numeracy Screening-Pilot

1	17	11	7
4	6	3	14

UNO PAM-RTI/PBS ©2009

10 8	8 2
14 5	1 3

STOP

UNO PAM-RTI/PBS ©2009

1	17	4	5
11	7	4	9
4	6	5	10
7	9	0	17
3	5	3	7

UNO PAM-RTI/PBS ©2009

15 14	2 4
8 0	3 14
2 17	10 9
10 8	5 9
7 9	8 2

UNO PAM-RTI/PBS ©2009

2 5	10 2
11 13	2 5
7 20	4 1
2 5	3 6
14 5	17 0

UNO PAM-RTI/PBS ©2009

1 4	13 5
15 17	2 3
5 2	20 15
1 4	16 6
8 2	0 10

UNO PAM-RTI/PBS ©2009

8 6	1 3
18 12	17 9
0 6	9 7
7 17	6 5
8 7	6 13

UNO PAM-RTI/PBS ©2009

6 16	9 4
8 10	12 7
5 4	13 8
0 6	13 18
1 13	8 7

UNO PAM-RTI/PBS ©2009

7	16

6	3

5	6

UNO PAM-RTI/PBS ©2009

Appendix O

Quantity Discrimination Probe 2

PAM Early Numeracy Screening-Pilot

5	3	1	10
4	9	6	18
0	2	11	2
2	4	20	11
6	1	2	10

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3	13	5	8
7	5	7	2
15	2	9	7
9	2	1	0
9	19	16	9

11 10	7 5
1 5	4 9
3 2	17 2
9 5	11 9
16 9	18 9

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

7	12	20	7
8	16	11	15
3	12	5	6
7	20	12	16
5	7	4	3

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

7 5	7 8
14 13	4 5
8 15	2 4
7 4	15 8
6 0	9 7

3	9

6	4

1	3

6	4

8	5

7	9

3	9

7	10

19	11

6	2

1	9

7	9

3	2

Appendix P

Directions for Quantity Discrimination Probes

PAM Early Numeracy Screening-Pilot

Instructions for Quantity Discrimination – Group Administration

Administration: Group

Time: 2 minutes

Materials:

- Student copy of Early Numeracy Group Probe Booklet
- Direction booklet & extra copy of student booklet
- Pencils
- Timer

Directions:

1. Place a student copy in front of each student. Say to the students: **"Please open your booklet to the first page and put your finger on the smiley face. (2) "Turn to the Pink Page."**
2. Say these specific directions to the students: **"Look at the page in front of you. In each row there are boxes with numbers in them."** (point to the first row of boxes on a copy which is shown to students.) **"Look at the first box, with the numbers 1 and 17. Pick-up your pencil and put a circle around the number that is greater."**
3. Check each student's response. If the response is correct: **"The number that is greater is 17. You should have circled the 17 because 17 is greater than 1."**
4. Continue with the other example(s) using the same specific directions.
5. Say to the students: **"Find the arrow on the bottom, this arrow means we need to turn the page and keep working. I want everyone to turn the page.'**
6. Now say these specific directions: "**Let's try the next one.** (Hold up the page and point to the next problem.) **Look at the box with the numbers 10 and 8. Put a circle around the number that is greater.** Walk around and check students' responses. **"The number that is greater is 10. You should have circled the 10 because 10 is greater than 8."** Continue with the other example(s) on the page using the same specific directions.
7. Now say to the students: **"Find the stop sign at the bottom of the page. This stop sign means we need to stop and put our pencils down** (Make sure all pencils are down).
8. Say to the students: **"Turn the page and put your finger on the smiley face. This page is the same as the one you just did. When I say begin, I want you to circle the number that is greater in each set of boxes. Start at the smiley face and go across the page** (demonstrate by pointing)**. Try each one. When you come to arrow at the end of the page, turn the page and continue working until you get to the stop sign. Are there any questions? Pick up your pencil. Put your finger on the first one. Ready, begin."**
9. At the end of 2 minutes, say, **"Stop."**

Scoring Rules:
Rule 1: If a student correctly identifies the number, score the item as correct.
Rule 2: If a student skips a problem, score the problem as incorrect.
Rule 3: If a student skips an entire row, mark each problem in the row as incorrect.
Rule 4: Count the number of correct responses on each page and record it on the line at the top of every page. Next, add the totals at the top of every page and record it on the Early Numeracy Summary Score Sheet.

Appendix Q
PAM Early Numeracy Score Sheet-Pilot

Early Numeracy Summary Score Sheet

INDIVIDUAL MEASURES Date: ____________

	Score (# correct)	PAM Team Member Code
Number Identification		
Form 1		
Form 2		
Mean		
Quantity Discrimination		
Form 1		
Form 2		
Mean		
Missing Number		
Form 1		
Form 2		
Mean		

GROUP MEASURES Date: ____________

	Score (# correct)	PAM Team Member Code
Number Identification		
Form 1		
Form 2		
Mean		
Quantity Discrimination		
Form 1		
Form 2		
Mean		
Missing Number		
Form 1		
Form 2		

Appendix R

Missing Number Probe 1

PAM Early Numeracy Screening-Pilot

2   3   4   ___ | 5   1   6 |

---

0   ___   2   3 | 4   2   1 |

---

10   20   30   ___ | 40   35   31 |

Example Sp1 - 1   OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

2	3	___	5		4	1	6
1	___	3	4		5	2	0
6	7	8	___		10	9	5

5	__	7	8	9 6 4
__	2	3	4	5 0 1
1	2	__	4	5 0 3
4	__	6	7	5 8 3
2	__	4	5	3 6 1
3	4	5	__	7 2 6

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

1 2 3 __	5	4	0
2 3 __ 5	4	1	6
60 70 80 __	85	81	90
__ 1 2 3	4	2	0
3 4 5 __	2	6	7
__ 7 8 9	6	10	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

20	30	40	___	45	50	41
1	___	3	4	5	2	0
1	2	3	___	4	0	5
4	___	6	7	5	8	3
___	4	5	6	7	2	3
5	6	7	___	8	9	4

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

__	4	5	6	7	3	2
7	8	__	10	11	6	9
20	25	30	__	35	40	31
__	2	3	4	5	1	0
3	4	__	6	7	5	2
6	7	8	__	10	9	5

5	10	15	__	16	20	25
0	1	2	__	1	3	4
7	__	9	10	8	6	11
25	30	35	__	40	45	36
2	3	4	__	5	6	1
1	2	3	__	0	4	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

___	7	8	9		5	10	6
20	30	40	___		50	41	45
4	___	6	7		3	8	5
0	___	2	3		2	4	1
3	___	5	6		2	7	4
1	2	3	___		5	0	4

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

10	15	20	___		30	25	21
35	40	45	___		46	55	50
1	___	3	4		2	5	0
1	2	3	___		5	0	4
___	3	4	5		1	6	2
___	6	7	8		4	5	9

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

6	__	8	9	10	7	5
7	8	9	__	10	11	6
7	8	__	10	9	11	6
5	6	7	__	9	4	8
__	3	4	5	1	6	2
__	7	8	9	6	5	10

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

___ 5 6 7	4	3	8
7 8 9 ___	6	11	10
15 20 25 ___	35	30	26
1 ___ 3 4	5	2	0
6 7 8 ___	10	9	5
0 1 2 ___	4	3	1

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

40	50	60	___	65	61	70
20	30	40	___	45	41	50
6	7	8	___	9	5	10
1	2	3	___	4	5	0
7	8	9	___	10	6	11
4	5	___	7	3	6	8

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3	4	5	___	2	6	7
5	6	7	___	4	9	8
6	7	8	___	9	10	5

Missing Number Probe 2

PAM Early Numeracy Screening-Pilot

0	1	2	__	3	4	1
3	4	5	__	2	6	7
__	5	6	7	3	8	4
0	1	2	__	4	3	1
__	4	5	6	7	2	3
2	3	4	__	5	6	1

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

5	6	7	__		9	8	4
6	7	8	__		5	9	10
__	6	7	8		5	4	9
7	__	9	10		6	8	11
__	4	5	6		2	3	7
5	10	15	__		25	20	16

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

20	25	30	__		35	31	40
7	8	__	10		11	6	9
6	7	8	__		10	9	5
7	8	9	__		10	11	6
70	80	90	__		91	100	95
30	40	50	__		55	60	51

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

1	__	3	4	0	5	2
10	15	20	__	25	30	21
0	__	2	3	4	1	2
4	5	__	7	3	6	8
3	4	5	__	2	7	6
5	6	7	__	9	4	8

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

__ 6 7 8	5	4	9
__ 3 4 5	1	2	6
2 3 4 __	5	1	6
30 40 50 __	60	51	55
1 2 3 __	0	5	4
4 __ 6 7	5	8	3

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

5	6	7	__	8	4	9
5	__	7	8	6	4	9
6	7	8	__	10	9	5
__	1	2	3	0	2	4
__	8	9	10	7	11	6
__	3	4	5	6	1	2

4	5	6	__	7	8	3
__	4	5	6	7	2	3
1	__	3	4	0	5	2
7	8	9	__	11	10	6
70	80	90	__	100	95	91
60	70	80	__	81	90	85

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

25	30	35	__	45	36	40
4	5	6	__	3	8	7
6	__	8	9	7	5	10
2	3	4	__	6	5	1
7	8	__	10	11	9	6
__	1	2	3	2	4	0

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3	4	5	__	6	7	2
2	__	4	5	6	3	1
10	20	30	__	40	31	35
2	__	4	5	3	1	6
0	1	2	__	1	3	4
7	8	9	__	10	6	11

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

0	1	2	__	3	4	1
3	4	5	__	6	7	2
6	7	8	__	9	10	5
2	3	__	5	4	6	1
__	3	4	5	1	2	6
25	30	35	__	45	36	40

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

1		3	4	2	5	0
6		8	9	10	5	7
2	3	4		1	5	6

Appendix T

Directions for Missing Number Probes

PAM Early Numeracy Screening-Pilot

Instructions for Missing Numbers Group Administration

Administration: Group

Time: 2 minutes

Materials:

- Student copy of Early Numeracy Group Probe Booklet
- Direction booklet & extra copy of student booklet
- Pencils
- Timer

Directions:

1. Say these specific directions to the students: (1) "**Turn to the Orange Page." (2)** "**Turn to the Yellow Page."**
2. Walk around the room and check to ensure every student book is turned to the Orange Page/Yellow Page.
3. Say these specific directions to the students: **"Turn the page and put your finger on the smiley face. Look at the row under the smiley face. It has the numbers 2, 3, 4. Now look at the box next to it. It has the numbers 5, 1, 6. Find the number that comes next in the pattern, pick up your pencil and circle it."** Walk around and check each student's response.
4. Now say these specific directions: **"The number that goes in the blank is 5. You should have circled the 5 because 5 comes next in the pattern (2,3,4, 5)."** Continue with the other example(s) on the page using the same specific directions.
5. Say to the students: **"Find the arrow on the bottom, this arrow means we need to turn the page and keep working. I want everyone to turn the page."**
6. Now say these specific directions: "**Let's try the next one."** (Hold up the page and point to the next problem.) **"Look at the row with numbers 2, 3, _, 5. Now look at the box next to it. It has the numbers 4, 1, 6. Find the number that comes next in the pattern and circle it."** Walk around and check students' responses. **"The number that goes in the blank is 4. You should have circled the 4 because 4 comes next in the pattern (2,3,4, 5)."** Continue with the other example(s) on the page using the same specific directions.
7. Now say to the students: **"Find the stop sign at the bottom of the page. This stop sign means we need to stop and put our pencils down"** (Make sure all pencils are down).
8. Say to the students: "**Turn the page. This page is the same as the one you just did. When I say, begin, I want you to circle the number that goes in each blank. Start at the top of the page and work down** (demonstrate by pointing)**. Try each one. Remember when you see an arrow to turn the page and keep working until you get to the stop sign. Are there any questions? Put your finger on the first one. Pick up your pencils. Ready, begin.** (start stopwatch)
9. At the end of 2 minutes, say, **"Stop. Pencils down."**

Scoring Rules:
Rule 1: If a student correctly circles /identifies the number, score the item as correct.
Rule 2: If a student skips a problem, score the problem as incorrect.
Rule 3: If a student skips an entire row, mark each problem in the row as incorrect.
Rule 4: Count the number of correct responses and record it on the Early Numeracy Summary Score Sheet

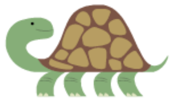
Appendix U
Number Identification Probe 1
PAM Early Numeracy Screening-Pilot

0	8	29	2

20	27	16	8

11	22	2	29

10	19	1	3

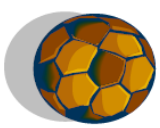


3	23	0	26

22	30	5	7

7	1	5	30

25	14	15	9



9	2	25	7

16	5	7	0

0	8	29	2

27	11	13	14

20	27	16	8

19	30	9	12

17	12	16	21

0	19	7	28

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

24	0	7	18

14	17	13	18

28	27	6	26

24	6	14	25

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix V
Number Identification Probe 2
PAM Early Numeracy Screening-Pilot

9	23	21	2

11	12	30	18

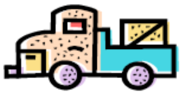
15	26	6	10

29	23	18	28

20	3	15	8

1	20	19	25

UNO PAM-RTI/PBS ©2009



13	4	14	5

11	26	4	25

29	15	5	11

15	24	16	28

Appendix W

Directions for Number Identification Probe 1

PAM Early Numeracy Screening-Pilot

Instructions for Number Identification Group Administration

(2 minutes, 20 seconds)

Say these specific directions to the students: **"Turn to the green page."** Walk around the room and check to ensure every student book is turned to the green page. Now say these specific directions to the students: **"Turn the page. Look at the page in front of you. Each row has a picture next to a box with numbers. Put your finger on the dog."** (Point to the first row of boxes on a copy which is shown to students.) Now the say the directions under each set of numbers.

0	8	29	2

"Put your finger on the dog and circle the 8." Allow the students 6 seconds to circle the number. Then hold-up a student booklet, point to the number 8 and say **"This is the number 8. You should have circled this number."** Pause and say, **"Let's try the next one."**

20	27	16	8

Say, **"Put your finger on the ice cream and circle the 16."** Allow the students 6 seconds to circle the number. Then hold-up a student booklet, point to the number 16 and say **"This is the number 16. You should have circled this number."** Pause and say, **"Let's try the next one. Turn the page."**

11	22	2	29

"Put your finger on the hat and circle the 22." Allow the students 6 seconds to circle the number. Then hold-up a student booklet, point to the number 22 and say **"This is the number 22. You should have circled this number."** Pause and say, **"Let's try the next one."**

10	19	1	3

"Put your finger on the dog and circle the 19." Allowthe students 6 seconds to circle the number. Then hold-up a student booklet, point to the number 19 and say **"This is the number 19. You should have circled this number."** Pause and say **"Let's try the next one. Turn the page and put your finger on the turtle."**

0s

Put your finger on the turtle and circle the 3

8s

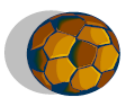
Put your finger on the rabbit and circle the 30

16s

Put your finger on the car and circle the 7

24s

Put your finger on the sun and circle the 9



32s

Put your finger on the ball and circle the 25

40s

Put your finger on the pencil and circle the 16

When the timer reaches **48 seconds** say,

"TURN THE PAGE and put your finger on the dog."

(Ten seconds are allocated to turn the page.)

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

58s

Put your finger on the dog and circle the 29

66s (1.06 min.)

Put your finger on the hat and circle the 11

74s (1.14 min.)

Put your finger on the snake and circle the 20

82s (1.22 min.)

Put your finger on the face and circle the 12

90s (1.30 min.)

Put your finger on the monkey and circle the 16

98s (1.38 min.)

Put your finger on the apple and circle the 28

When the timer reaches **106 seconds (1.46 min.)** say,

"TURN THE PAGE and put your finger on the truck."

(Ten seconds are allocated to turn the page.)

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

116s (1.56 min)

Put your finger on the truck and circle the 0

124s (2.04 min.)

Put your finger on the star and circle the 13

132s (2.12 min.)

Put your finger on the tree and circle the 28

When the timer reaches **140 seconds (2.20 min.)** SAY,

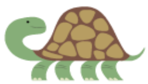
"STOP. PENCILS DOWN."

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix X

Directions for Number Identification Probe 2

PAM Early Numeracy Screening-Pilot



0s

Put your finger on the turtle and circle the 0

8s

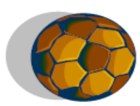
Put your finger on the rabbit and circle the 9

16s

Put your finger on the car and circle the 28

24s

Put your finger on the sun and circle the 7



32s

Put your finger on the ball and circle the 8

40s

Put your finger on the pencil and circle the 11

When the timer reaches **48 seconds** say,

"TURN THE PAGE and put your finger on the dog."

(Ten seconds are allocated to turn the page.)

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

58s

Put your finger on the dog and circle the 2

66s (1.06 min.)

Put your finger on the hat and circle the 30

74s (1.14 min.)

Put your finger on the snake and circle the 6

82s (1.22 min.)

Put your finger on the face and circle the 18

90s (1.30 min.)

Put your finger on the monkey and circle the 20

98s (1.38 min.)

Put your finger on the apple and circle the 19

When the timer reaches **106 seconds (1.46 min.)** say,
"TURN THE PAGE and put your finger on the truck."
(Ten seconds are allocated to turn the page.)

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

116s (1.56 min)

Put your finger on the truck and circle the 13

124s (2.04 min.)

Put your finger on the star and circle the 11

132s (2.12 min.)

Put your finger on the tree and circle the 5

When the timer reaches **140 seconds (2.20 min.)** SAY,

"STOP. PENCILS DOWN."

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix Y

Parent Consent for Pilot Study

UNIVERSITY of NEW ORLEANS

DEPARTMENT OF SPECIAL EDUCATION
AND HABILITATIVE SERVICES

Dear Parent:

I am a professor College of Education and Human Development at the University of New Orleans. I am conducting a research study to investigate the possibility of using group math probes with Kindergarten students. Currently these probes are given individually, however time would be saved if they could be administered in a group setting if the results are comparable.

I am requesting your child's participation, which will involve administering up to three one page math problems which would take about 1 minute each. Yous child already receives probes given on an individual basis, so this is not an unfamiliar activity. Your child's participation in this study is voluntary. If you choose not to have your child participate or to withdraw your child from the study at any time, there will be no penalty, and it will not affect your child's grade in any way. Likewise, if your child chooses not to participate or to withdraw from the study at any time, there will be no penalty. The results of the research study may be published, but your child's name will not be used.

Although there may be no direct benefit to your child, the possible benefit of your child's participation is the saving of time for both the teacher and the student in administering these basic math probes.

The risks associated with participating are minimal are not greater than those ordinarily encountered during the performance of routine academic tasks.

If you have any questions concerning the research study or your child's participation in this study, please call me at (504) 280-5688, or contact your child's teacher.

If you have any questions about you or your child's rights as a subject/participant in this research, or if you feel you or your child have been placed at risk, you can contact Dr. Ann O'Hanlon at the University of New Orleans at 504-280-6501.

Sincerely,

Gale M. Naquin, Ph.D.

By signing below, you are giving consent for your child _______________ to participate in the above study.

_____________________ _____________________ _____

Signature Printed Name Date

246 Bicentennial Education Center Lakefront Campus 2000 Lakeshore Drive New Orleans, Louisiana 70148 504.280.6609 Fax 504.280.5588
A Member of the Louisiana State University System Committed to Equal Opportunity

Appendix Z
Early Numeracy Fidelity Checklist for
Curriculum-Based Measures

Early Numeracy Fidelity Checklist

PAM Code: ____________________ DATE: __________________
Observer: ______________________

Key: - Incorrect

+ Correct

Testing Procedure	Probe	Probe
1. Has materials needed		
2. Reads the standardized directions to the student		
3. Says "Begin"		
4. Starts timer		
5. Monitors students taking the test by walking around the room checking to make sure they are completing each page and turning the page **(group measures)**		
6. Delivers probe items at an appropriate rate **(individual measures)**		
7. Times accurately		
8. Says "stop"		
9. Correctly scores probes		
10. Correctly records score on score summary sheet		

Appendix AA
Early Numeracy Fidelity Checklist for
Standardized Tests

Early Numeracy Fidelity Checklist

Examiner Code: ____________________ DATE: ___________________

Observer: _________________________

Key: - Incorrect

+ Correct

Testing Procedure	Assessment	Assessment
1. Has materials needed		
2. Reads the standardized directions to the student		
3. Provides appropriate feedback to the examinee		
4. Establishes a Basel		
5. Establishes a Ceiling		
6. Correctly scores items		
7. Correctly records score on protocol		
8. Correctly converts raw scores		
9. Other		

Appendix BB

Modifications

Quantity Discrimination

PAM Early Numeracy Screening

____

5	3	1	10
4	9	6	18
0	2	11	2
2	4	20	11
6	1	2	10

Appendix CC

PAM Early Numeracy Summary Score Sheet

Date: ______________________

PAM Early Numeracy Summary Score Sheet

Group Probes

	# of correct responses	**Scorer ID Code**
Quantity Discrimination		
Probe 1		
Probe 2		
Mean		
Missing Number		
Probe 1		
Probe 2		
Mean		
Number Identification		
Probe 1		
Probe 2		
Mean		

Individual Probes

Number Identification	**# of correct responses**	**Evaluator ID Code**	**Scorer ID Code**
Probe 1			
Probe 2			
Mean			

Appendix DD

Quantity Discrimination Probe 1

PAM Early Numeracy Screening

1	17	11	7
4	6	3	14

10 8	8 2
14 5	1 3

1	17	4	5
11	7	4	9
4	6	5	10
7	9	0	17
3	5	3	7

15 14	2 4
8 0	3 14
2 17	10 9
10 8	5 9
7 9	8 2

2 5	10 2
11 13	2 5
7 20	4 1
2 5	3 6
14 5	17 0

1 4	13 5
15 17	2 3
5 2	20 15
1 4	16 6
8 2	0 10

8 6	1 3
18 12	17 9
0 6	9 7
7 17	6 5
8 7	6 13

6 16	9 4
8 10	12 7
5 4	13 8
0 6	13 18
1 13	8 7

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

7 16

6 3

5 6

5	3	1	10
4	9	6	18
0	2	11	2
2	4	20	11
6	1	2	10

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3 13	5 8
7 5	7 2
15 2	9 7
9 2	1 0
9 19	16 9

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

11 10	7 5
1 5	4 9
3 2	17 2
9 5	11 9
16 9	18 9

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

7 12	20 7
8 16	11 15
3 12	5 6
7 20	12 16
5 7	4 3

7	5	7	8
14	13	4	5
8	15	2	4
7	4	15	8
6	0	9	7

3 9	6 4
1 3	6 4
8 5	7 9
3 9	7 10
19 11	6 2

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

1 9	7 9
3 2	

Appendix FF
Directions for Quantity Discrimination Probes
PAM Early Numeracy Screening

Directions for Quantity Discrimination Group Administration

1. Place a student booklet in front of each student.

   "Open your booklet to the first page and put your finger on the smiley face."

2. Wait for the students to put their finger on the smiley face. (*Point to the first row of boxes on the backside of the page.)*

   "Look at the page in front of you. In each row there are boxes with numbers."

   "Look at the first box with the numbers 1 and 17. Pick up your pencil and circle the number that is greater."

3. Walk around to check the students have answered the appropriate examples.

   "The number that is greater is 17. You should have circled 17, because 17 is greater than 1. Let's try another one. (*Point to the box with 11 and 7*) **Look at the next box, with the numbers 11 and 7. Circle the number that is greater."**

4. Walk around to check the students have answered the appropriate examples.

   "The number that is greater is 11. You should have circled 11, because 11 is greater than 7. (*Point to the box with 4 and 6*) **Find the box with the numbers 4 and 6. Circle the number that is greater."**

5. Walk around to check the students have answered the appropriate examples.

   "The number that is greater is 6. You should have circled 6, because 6 is greater than 4. Let's try another one. (*Point to the box with 3 and 14*) **Look at the next box, with the numbers 3 and 14. Circle the number that is greater."**

6. Walk around to check the students have answered the appropriate examples.

   "The number that is greater is 14. You should have circled 14, because 14 is greater than 3. Find the arrow on the bottom of the page. This arrow means we need to turn the page and keep working. I want everyone to turn the page."

7. Place a student booklet in front of each student.

   "Open your booklet to the first page and put your finger on the smiley face."

8. Wait for the students to put their finger on the smiley face. (*Point to the first row of boxes on the backside of the page.)*

   "Look at the page in front of you. In each row there are boxes with numbers."

   "Look at the first box with the numbers 1 and 17. Pick up your pencil and circle the number that is greater."

9. Walk around to check the students have answered the appropriate examples.

   "The number that is greater is 17. You should have circled 17, because 17 is greater than 1. Let's try another one. (*Point to the box with 11 and 7*) **Look at the next box, with the numbers 11 and 7. Circle the number that is greater."**

10. Walk around to check the students have answered the appropriate examples.

    "The number that is greater is 11. You should have circled 11, because 11 is greater than 7. (*Point to the box with 4 and 6*) **Find the box with the numbers 4 and 6. Circle the number that is greater."**

11. Walk around to check the students have answered the appropriate examples.

    "The number that is greater is 6. You should have circled 6, because 6 is greater than 4. Let's try another one. (*Point to the box with 3 and 14*) **Look at the next box, with the numbers 3 and 14. Circle the number that is greater."**

12. Walk around to check the students have answered the appropriate examples.

    "The number that is greater is 14. You should have circled 14, because 14 is greater than 3. Find the arrow on the bottom of the page. This arrow means we need to turn the page and keep working. I want everyone to turn the page."

Appendix GG

Missing Number Probe 1

PAM Early Numeracy Screening

2 3 4 __ | 5 1 6

0 __ 2 3 | 4 2 1

10 20 30 __ | 40 35 31

Example MN 1 - 1 OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

2	3	___	5	4	1	6

1	___	3	4	5	2	0

6	7	8	___	10	9	5

___

5	___	7	8	9	6	4
___	2	3	4	5	0	1
1	2	___	4	5	0	3
4	___	6	7	5	8	3
2	___	4	5	3	6	1
3	4	5	___	7	2	6

______

				Choices		
1	2	3	___	5	4	0
2	3	___	5	4	1	6
60	70	80	___	85	81	90
___	1	2	3	4	2	0
3	4	5	___	2	6	7
___	7	8	9	6	10	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

20	30	40	__	45	50	41
1	__	3	4	5	2	0
1	2	3	__	4	0	5
4	__	6	7	5	8	3
__	4	5	6	7	2	3
5	6	7	__	8	9	4

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

___	4	5	6	7	3	2
7	8	___	10	11	6	9
20	25	30	___	35	40	31
___	2	3	4	5	1	0
3	4	___	6	7	5	2
6	7	8	___	10	9	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

5	10	15	__	16	20	25
0	1	2	__	1	3	4
7	__	9	10	8	6	11
25	30	35	__	40	45	36
2	3	4	__	5	6	1
1	2	3	__	0	4	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

__	7	8	9	5	10	6
20	30	40	__	50	41	45
4	__	6	7	3	8	5
0	__	2	3	2	4	1
3	__	5	6	2	7	4
1	2	3	__	5	0	4

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

10	15	20	___	30	25	21
35	40	45	___	46	55	50
1	___	3	4	2	5	0
1	2	3	___	5	0	4
___	3	4	5	1	6	2
___	6	7	8	4	5	9

___

6	___	8	9	10	7	5
7	8	9	___	10	11	6
7	8	___	10	9	11	6
5	6	7	___	9	4	8
___	3	4	5	1	6	2
___	7	8	9	6	5	10

__	5	6	7	4 3 8
7	8	9	__	6 11 10
15	20	25	__	35 30 26
1	__	3	4	5 2 0
6	7	8	__	10 9 5
0	1	2	__	4 3 1

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

40	50	60	__	65	61	70
20	30	40	__	45	41	50
6	7	8	__	9	5	10
1	2	3	__	4	5	0
7	8	9	__	10	6	11
4	5	__	7	3	6	8

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3	4	5	___	2	6	7
5	6	7	___	4	9	8
6	7	8	___	9	10	5

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

Appendix HH
Missing Number Probe 2
PAM Early Numeracy Screening

______

0	1	2	__	3	4	1
3	4	5	__	2	6	7
__	5	6	7	3	8	4
0	1	2	__	4	3	1
__	4	5	6	7	2	3
2	3	4	__	5	6	1

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

5	6	7	__	9	8	4
6	7	8	__	5	9	10
__	6	7	8	5	4	9
7	__	9	10	6	8	11
__	4	5	6	2	3	7
5	10	15	__	25	20	16

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

20	25	30	__	35	31	40
7	8	__	10	11	6	9
6	7	8	__	10	9	5
7	8	9	__	10	11	6
70	80	90	__	91	100	95
30	40	50	__	55	60	51

______

1	___	3	4		0	5	2
10	15	20	___		25	30	21
0	___	2	3		4	1	2
4	5	___	7		3	6	8
3	4	5	___		2	7	6
5	6	7	___		9	4	8

______

__	6	7	8	5	4	9
__	3	4	5	1	2	6
2	3	4	__	5	1	6
30	40	50	__	60	51	55
1	2	3	__	0	5	4
4	__	6	7	5	8	3

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

5	6	7	__	8	4	9
5	__	7	8	6	4	9
6	7	8	__	10	9	5
__	1	2	3	0	2	4
__	8	9	10	7	11	6
__	3	4	5	6	1	2

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

4	5	6	__	7	8	3
__	4	5	6	7	2	3
1	__	3	4	0	5	2
7	8	9	__	11	10	6
70	80	90	__	100	95	91
60	70	80	__	81	90	85

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

25	30	35	__	45	36	40
4	5	6	__	3	8	7
6	__	8	9	7	5	10
2	3	4	__	6	5	1
7	8	__	10	11	9	6
__	1	2	3	2	4	0

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

3	4	5	___		6	7	2
2	___	4	5		6	3	1
10	20	30	___		40	31	35
2	___	4	5		3	1	6
0	1	2	___		1	3	4
7	8	9	___		10	6	11

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

______

0	1	2	__	3	4	1
3	4	5	__	6	7	2
6	7	8	__	9	10	5
2	3	__	5	4	6	1
__	3	4	5	1	2	6
25	30	35	__	45	36	40

OSEP Award #H324H030003 modified by UNO PAM-RTI/PBS 2009

1 3 4 | 2 5 0

6 8 9 | 10 5 7

2 3 4 | 1 5 6

STOP

Appendix II
Directions for Missing Number Probes
PAM Early Numeracy Screening

Instructions for Missing Numbers Group Administration

Administration: Group

Time: 2 minutes

Materials:

- Student copy of Early Numeracy Group Probe Booklet
- Direction booklet & extra copy of student booklet
- Pencils
- Timer

Directions:

1. Say these specific directions to the students: (1) "**Turn to the Orange Page." (2)** "**Turn to the Yellow Page."**
2. Walk around the room and check to ensure every student book is turned to the Orange Page/Yellow Page.
3. Say these specific directions to the students: **"Turn the page and put your finger on the smiley face. Look at the row under the smiley face. It has the numbers 2, 3, 4. Now look at the box next to it. It has the numbers 5, 1, 6. Find the number that comes next in the pattern, pick up your pencil and circle it."** Walk around and check each student's response.
4. Now say these specific directions: **"The number that goes in the blank is 5. You should have circled the 5 because 5 comes next in the pattern (2,3,4, 5)."** Continue with the other example(s) on the page using the same specific directions.
5. Say to the students: **"Find the arrow on the bottom, this arrow means we need to turn the page and keep working. I want everyone to turn the page."**
6. Now say these specific directions: "**Let's try the next one."** (Hold up the page and point to the next problem.) **"Look at the row with numbers 2, 3, _, 5. Now look at the box next to it. It has the numbers 4, 1, 6. Find the number that comes next in the pattern and circle it."** Walk around and check students' responses. **"The number that goes in the blank is 4. You should have circled the 4 because 4 comes next in the pattern (2,3,4, 5)."** Continue with the other example(s) on the page using the same specific directions.
7. Now say to the students: **"Find the stop sign at the bottom of the page. This stop sign means we need to stop and put our pencils down"** (Make sure all pencils are down).
8. Say to the students: "**Turn the page. This page is the same as the one you just did. When I say, begin, I want you to circle the number that goes in each blank. Start at the top of the page and work down** (demonstrate by pointing)**. Try each one. Remember when you see an arrow to turn the page and keep working until you get to the stop sign. Are there any questions? Put your finger on the first one. Pick up your pencils. Ready, begin.** (start stopwatch)
9. At the end of 2 minutes, say, **"Stop. Pencils down."**

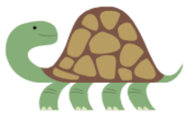
Scoring Rules:
Rule 1: If a student correctly circles /identifies the number, score the item as correct.
Rule 2: If a student skips a problem, score the problem as incorrect.
Rule 3: If a student skips an entire row, mark each problem in the row as incorrect.
Rule 4: Count the number of correct responses and record it on the Early Numeracy Summary Score Sheet

Appendix JJ

Number Identification Probe 1

PAM Early Numeracy Screening

____

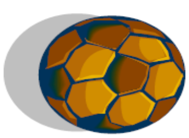


14	10	**0**	20

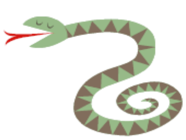
16	20	26	6

2	12	22	7

19	29	10	9



8	18	7	28

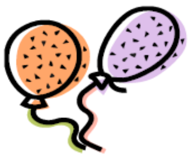


10	15	20	30

	3	23	13	0
	27	16	17	7
	30	5	20	10
	25	15	3	5
	24	14	4	9
	25	15	11	5

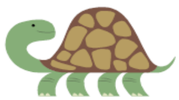
	1	11	20	21
	26	12	16	6
	22	6	2	12
	2	18	8	28
	17	23	13	3
	27	7	17	11

____

	21	14	11	1
	9	19	4	29
	14	2	24	4

Appendix KK
Number Identification Probe 2
PAM Early Numeracy Screening

____

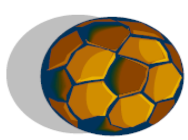


16	26	18	6

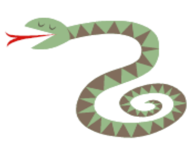
13	0	10	20

19	29	0	9

18	8	28	3



17	22	2	12



30	20	5	10

______

10	15	30	20

13	5	15	25

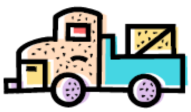
3	10	13	23

24	14	4	19

17	7	6	27

25	15	1	5

26	2	16	6



22	2	16	12

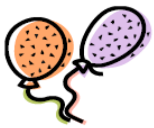
7	23	13	3

1	11	20	21

27	7	17	11

12	8	18	28

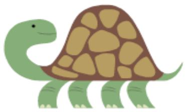
4	21	11	1



9	14	19	29

14	12	24	4

Appendix LL
Directions for Number Identification Probe 1
PAM Early Numeracy Screening


0s

Put your finger on the turtle, circle the **0**


5s

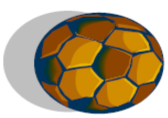
Put your finger on the rabbit, circle the **6**

10s

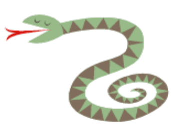
Put your finger on the car, circle the **2**


15s

Put your finger on the sun, circle the **19**


20s

Put your finger on the ball, circle the **18**


25s

Put your finger on the snake, circle the **10**

When the timer reaches 30 seconds say,
"Turn the page and put your finger on the dog"
(Ten seconds are allocated to turn the page.)

40s

Put your finger on the dog, circle the **13**

45s

Put your finger on the hat, circle the **7**

50s

Put your finger on the ice cream, circle the **20**

55s

Put your finger on the smiley face, circle the **15**

60s

Put your finger on the monkey, circle the **14**

65s

Put your finger on the apple, circle the **5**

When the timer reaches 70 seconds (1.10 min.) say,
"Turn the page and put your finger on the house."
(Ten seconds are allocated to turn the page.)

80s (1.20min)

Put your finger on the house, circle the **1**

85s (1.25 min)

Put your finger on the truck, circle the **16**

90s (1.30 min)

Put your finger on the flower, circle the **12**

95s (1.35 min)

Put your finger on the swing, circle the **8**

100s (1.40min)

Put your finger on the tree, circle the **3**

105s (1.45min)

Put your finger on the star, circle the **17**

When the timer reaches 110 seconds (1.55 min.) say,
"Turn the page and put your finger on the cat."
(Ten Seconds are allocated to turn the page.)

120s (2.00 min)

Put your finger on the cat, circle the **11**

125s (2.05 min)

Put your finger on the balloons, circle the **9**

130s (2.10 min)

Put your finger on the pencil, circle the **4**

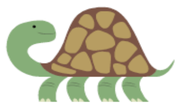
When the timer reaches 140 seconds (2.15 min.)

SAY, "STOP. PENCILS DOWN."

Appendix MM

Directions of Number Identification Probe 2

PAM Early Numeracy Screening



0s

Put your finger on the turtle, circle the **6**

5s

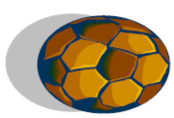
Put your finger on the rabbit, circle the **0**

10s

Put your finger on the car, circle the **19**

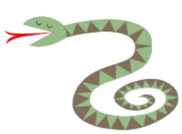
15s

Put your finger on the sun, circle the **18**



20s

Put your finger on the ball, circle the **2**



25s

Put your finger on the snake, circle the **10**

When the timer reaches 30 seconds say,

"Turn the page and put your finger on the dog"

(Ten seconds are allocated to turn the page.)

40s

Put your finger on the <u>dog</u>, circle the **20**

45s

Put your finger on the <u>hat</u>, circle the **15**

50s

Put your finger on the <u>ice cream</u>, circle the **13**

55s

Put your finger on the <u>smiley face</u>,circle the **14**

60s

Put your finger on the <u>monkey</u>, circle the **7**

65s

Put your finger on the <u>apple</u>, circle the **5**

When the timer reaches 70 seconds (1.10 min.) say,
"Turn the page and put your finger on the house."
(Ten seconds are allocated to turn the page.)

80s (1.20min)

Put your finger on the house, circle the **16**

85s (1.25 min)

Put your finger on the truck, circle the **12**

90s (1.30 min)

Put your finger on the flower, circle the **3**

95s (1.35 min)

Put your finger on the swing, circle the **1**

100s (1.40min)

Put your finger on the tree, circle the **17**

105s (1.45min)

Put your finger on the star, circle the **8**

When the timer reaches 110 seconds (1.55 min.) say,
"Turn the page and put your finger on the cat."
(Ten Seconds are allocated to turn the page.)

120s (2.00 min)

Put your finger on the cat, circle the **11**

125s (2.05 min)

Put your finger on the balloons, circle the **9**

130s (2.10 min)

Put your finger on the pencil, circle the **4**

When the timer reaches 135 seconds (2.15 min.)
SAY, "STOP. PENCILS DOWN."

Appendix NN
Early Math Measures Study Teacher Rating of
Students' Math Proficiency

Early Math Measures Study

Teacher Rating of Students' Math Proficiency

Teacher Name: ______________________________

Directions: Please list the names of each of the students participating in the project below. Think about each student in the context of peers of the same age/grade level. Please rate each student's ***general proficiency in math*** relative to other students in the same grade level. Students who have very low levels of math proficiency compared to their peers should be rated a 1. Those who have very high levels should be rated a 7. Thank you for your help!

Student Name	**1** *(least proficient)*	**2**	**3**	**4**	**5**	**6** *(most proficient)*	**7**
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7

Appendix OO

Parent Consent

UNIVERSITY of NEW ORLEANS

DEPARTMENT OF SPECIAL EDUCATION AND HABILITATIVE SERVICES

Dear Parent:

Our team, the Pupil Assistance Model Team, has been working with your parish for several years. We are researching the use of group administered math probes with kindergarten students. A math probe is a tool used to determine a student's knowledge in math. Usually these probes are given individually, however we will be administering probes in a group setting.

We are requesting your child's participation in this study which will involve your child completing three two-minute math probes in a group setting and a one-minute individual math probe. Your child will also be given an individually and group administered math assessment which will take no more than one hour. This is an activity your child will be familiar with and no harm from participating in the study is anticipated. If you choose not to have your child participate or to withdraw your child from the study at any time, there will be no penalty, and it will not affect your child's grade in any way. If your child chooses not to participate or to withdraw from the study at any time, there will be no penalty or affect on their grade. The results of the research study may be published, but your child's name will not be used.

There is not an immediate benefit for your child's participation in this study. However, the benefit of group administration may quickly and accurately identify students who need more help. Additionally, since group administration may save time, teachers may have more time for teaching and students may have more time for learning.

If you have any questions concerning the research study or your child's participation in this study, please call me at (504) 280-7221, or contact your child's teacher.

If you have any questions about you or your child's rights as a subject/participant in this research, or if you feel you or your child have been placed at risk, you can contact Dr. Ann O'Hanlon at the University of New Orleans at 504-280-6501.

Sincerely,

______________________	______________________	______________________
Stacy Winck, M.C.D.	Paulette J. Thomas, Ph.D.	Gale M. Naquin, Ph.D.
Doctoral Candidate	Professor of Special Education	Executive Director
Pupil Assistance Model	Pupil Assistance Model	Pupil Assistance Model
University of New Orleans	University of New Orleans	University of New Orleans

By signing below, you are giving consent for your child ___________________________ to participate in the above study.

______________________ ______________________ ________

Signature Printed Name Date

246 Bicentennial Education Center Lakefront Campus 2000 Lakeshore Drive
New Orleans, Louisiana 70148 504.280.6609 Fax 504.280.5588
A Member of the Louisiana State University System Committed to Equal Opportunity

Appendix PP
Request for IRB Change

THE UNIVERSITY *of* NEW ORLEANS

DEPARTMENT OF SPECIAL EDUCATION AND HABILITATIVE SERVICES

January 20, 2010

Robert Laird, Ph. D.
Institutional Review Board Chair
Associate Professor
Department of Psychology
Geology & Psychology Building
University of New Orleans
New Orleans, Louisiana 70148

Re: Request IRB change to 05Apr08

Dr. Laird:

Thank you for your assistance in the preparation of our request for the modification of IRB application 05Apr08. This research project was previously exempted by the IRB based on the exemptions listed in Category 1: Research conducted in established commonly accepted educational settings, involving normal educational practices. The research project was completed in Spring 2009, and we are currently requesting continuation of the project with modifications. None of the modifications change the status of the project or the risk to participate. We are requesting that Stacy Winck, UNO doctoral candidate, be added as a co-investigator to the project which will be her dissertation. Other modifications include collecting data using the *KeyMath-3 Diagnostic Assessment* and math subtests from the *Metropolitan Achievement Tests, 8th Edition*, and administering a revised version of the original mathematics probes. Please see attached revised parental consent form and mathematics probes. As previously stated these changes will not affect the status of the project or increase the risk to participate.

Thank you for consideration of this request. If we can be of further assistance or any additional information is required, please feel free to contact us at (504) 280-7221.

Sincerely,

Gale M. Naquin, Ph. D.
Executive Director
Pupil Assistance Model
Principal Investigator
Co-Chair, Dissertation Committee

Paulette J. Thomas, Ph.D.
Professor of Special Education
Major Professor
Principal Investigator
Pupil Assistance Model

246 Bicentennial Education Center *Lakefront Campus* 2000 Lakeshore Drive New Orleans, Louisiana 70148 504.280.6609 *fax* 504.280.5588

A Member of the Louisiana State University System Committed to Equal Opportunity

Appendix QQ
Human Subject Approval

University Committee for the Protection of Human Subjects in Research University of New Orleans

---

Campus Correspondence

Principal Investigator: Gale M. Naquin

Co-Investigator: Stacy Winck

Date: February 10, 2010

RE: "Investigating of Basic Group Mathematics Probes in Kindergarten Classrooms"

IRB#: 05Apr08

Your modification request was eligible for expedited review as the modifications did not change the potential risk to the participants. Modifications listed below have been approved.

- Stacy Winck was added as a co-investigator
- Modifications to collecting data using the *Key Math-3 Diagnostic Assessment* and math subsets from the *Metropolitan Achievement test 8th Edition* were added
- Mathematics probes were revised

Please remember that approval is only valid for one year from the approval date. Any changes to the procedures or protocols must be reviewed and approved by the IRB prior to implementation.

If an adverse, unforeseen event occurs (e.g., physical, social, or emotional harm), you are required to inform the IRB as soon as possible after the event.

Best of luck with your project!
Sincerely,

Robert Laird, Chair
UNO Committee for the Protection of Human Subjects in Research

VITA

The author was born in Jefferson Parish, Louisiana. She obtained her Bachelor's degree in education from University of New Orleans in1992, and her Master's degree in communication disorders from Louisiana State University Medical Center, School of Allied Health in 1994. For the last thirteen years, the author has provided diagnostic and intervention services to students with learning differences. She entered the University of New Orleans graduate program in Special Education to pursue a PhD in the area of tests and measurements.