NORTH CAROLINA PUBLIC KINDERGARTEN TEACHER PERCEPTIONS ON READINESS AND THE DEVELOPMENT OF A STATE K-3 ASSESSMENT

by

Chinmayi Sharma

A Thesis Presented to the

Sanford School of Public Policy

at

Duke University

May 2013

Abstract

An established pioneer in matters of early childhood education, North Carolina will be using the Race to the Top grant money they won in 2010 to create and implement a statewide K-3 assessment. The Department of Public Instruction formed a think tank to develop the vision for the tool and the final product will be administered to all public schools in 2015. This study is one facet of a multidimensional project aimed at bringing several key stakeholders to the table to discuss the development of this assessment instrument. In this project, teachers were asked to rate 22 items suggested for inclusion in the assessment on a 1-6 scale in terms of importance, 1 being the least important and 6 being the most important information for their day-to-day classroom instruction. Respondents were also asked to comment on the helpfulness and delivery of the assessment as well. Previous studies led to the hypothesis that teachers would prioritize the Social/Emotional Development domain of assessment over Cognitive Development or Literacy. The data were gathered by an online survey sent in an email to 9,493 public school kindergarten teachers and the respondents comprised a volunteer sample of 2,596 cases, a 25.62% response rate. Results showed that teachers placed greatest importance on Language and Literacy Development, followed by Cognitive Development. Parent's Education Level was given the lowest mean rate of importance out of the 22 suggested items. There was broad support for the assessment and the use of electronic tablets in its delivery. However, teachers did not support the option of the assessment being delivered through home visits. Demographic variations by region, locale size, education level and experience of teachers existed but were not robust.

Introduction

North Carolina has long seen itself as a leader in early childhood education and has professed a sincere dedication to the improvement of their education system as a whole. From Smart Start to More at Four—two unique early childhood education programs that have been expanded in North Carolina over the years—the state has already invested significant resources to close the achievement gap among various races and socio-economic groups upon entering Kindergarten. Many states in the U.S. already implement kindergarten entry assessments for their value as tracking devices of student progress and screening tools of potential concerns in a specific population or the system overall. North Carolina has long used school assessments, but now, has taken on the challenge of improving the assessments in the most cost-efficient way to be able to further improve the public education system.

The North Carolina Department of Public Instruction has received funding from the Race to the Top Early Learning Challenge grant and authorization from the General Assembly to develop an assessment that will generate a Child Profile at kindergarten entry and assess student progress through third grade. The state, under the orders of Superintendent June Atkinson, has formed a think tank of scholars and practitioners from various fields of education to psychology to coalesce their expertise and create a vision for kindergarten entry assessment in the most effective and efficient manner. Representatives hail from various parts of the state and act as the voices of different populations at the table. The think tank shall work until August of 2013 to create a vision for the assessment.

At that time, a design team will then spend the next year developing the assessment tool itself, finally implementing it on a statewide basis beginning in 2015.

This report is one part of a multi-step process of gathering information from various researchers and stakeholders to inform think tank members. This particular study is meant to gather the opinions of one particular stakeholder group: kindergarten teachers in state public schools. Their responses will reflect their evaluation of various assessment items in terms of importance to their day-to-day work. As the group that will be administering this assessment and stands to gain from its results, its opinions are highly relevant. As with all policy changes, it is imperative that all relevant parties provide their perspectives.

Background of the project

In January 2010, North Carolina became one of only twelve recipients of a $400 million dollar grant to be used over four years, expiring in 2014 for their outstanding performance in the federal Race to the Top (RttT) competition. The utilization of the funds was to be overseen by the state Department of Public Instruction in collaboration with the State Board of Education and the Governor's Education Transformation Commission. The funds have been used in various ways to spur public school innovation and move the state towards greater levels of improvement in public education (N.C. Race to the Top).

Some of these funds were set aside for the creation of an assessment tool to close the school readiness gap at Kindergarten entry and track students' progress through their early years in school to determine how best to instruct them. The funds are not only for the development of the assessment tool but also for the professional training required to prepare teachers to administer the assessment and state officials to analyze the results and be able to respond to them appropriately. In the application for RttT grant money, the state included a comprehensive history of childhood education assessments used up till that point and their proposed strategy to revamp the assessments to allow for an improved ability to meet students' needs. The initial plan was approved in 2010 with the allotment of funds and was revised July 5, 2012, to include changes in the plan as determined by the state (N.C. Race to the Top).

The changes were incorporated in the session laws passed in 2012, by the state legislature, with the urging from Democrat Governor Bev Perdue, to mandate the creation and efficient implementation of the assessments aforementioned in the state's RttT grant application. The first, "§115C-83.1E. Developmental screening and kindergarten entry assessment," was part of the Excellent Public Schools Act and called for the State Board of Education to ensure that all students are administered a developmental screening of early language, literacy and math skills within 30 days of their enrollment in school. Further, the state must ensure that each child completes a kindergarten entry assessment within 60 days of enrollment and said assessment must include the five essential domains of school readiness: language and literacy development, cognition and general knowledge, approaches toward learning, physical well-being and motor development and social and emotional development. The overall assessment must be aligned with the North Carolina standard of early learning and development as well as the standard course of study. Finally, the session law mandated that the assessment be administered at a classroom level and be able to account for children with various disabilities or non-English speakers. The timeline set by the session law dictates that while the law itself goes into effect in the 2013-2014 school year, the developmental screenings and kindergarten entry assessment are to be implemented at the beginning of the 2014-2015 school year (NCGA statutes). The second piece of relevant legislation, §115C-174.11 Components of the testing program, dictated that the State Board of Education is required to develop, adopt and provide developmentally appropriate individualized assessment instruments for a K-3 assessment. The results of these instruments shall be used to diagnose difficulties, propose remedies and be used for general progress analysis (NCGA statutes). The significance of this session law is that the onus of responsibility was put on the state to provide the resources for this project and facilitate its incorporation in North Carolina's 115 school districts.

The guiding principles of the creation of this tool will be the recommendations by the National Research Council's 2008 "Early Childhood Assessment: Why, What, and How?" report. The report's criteria place emphasis on the five domains of school readiness, the obligation to cater to children with high needs or disabilities, the incorporation of previous research and expert opinion in the creation of the tool and proper training of all involved parties both in the administration of the assessment and its evaluation. Finally, the report encourages a consistent sample testing of the assessments to ensure reliability but also warns extreme caution in the determination of the assessment's conclusions made on programs serving Children with High Needs or for accountability and high-stakes purposes (teacher evaluation or English Language Development programs).

The ultimate goal of the endeavor is to produce a formative assessment tool to be implemented at the start of Kindergarten, first-grade, second-grade and third-grade. The assessment will be a process used by teachers during instruction to provide feedback to adjust ongoing teaching, as well as to inform larger systematic changes in the field of public education.

After the administration of the initial assessment tool to the general kindergarten entry population, a "Child Profile" will be created for each student to address individualized challenges and track growth over the years. Currently, the only factor of readiness for kindergarten in North Carolina is that the child must be 5-years-old on or before August 31st of the year they wish to enter kindergarten. This surveying tool will not change this standard, but rather inform how best to minimize pre-existing gaps between children as they enter the school system. The state think tank has been keeping in mind the legislation and recommendations in the creation of this assessment. The timeline will be followed and the ultimate product will be put in full state use by June 2015.

History of kindergarten assessments in North Carolina

North Carolina currently has two sources of data to inform the Kindergarten Entry Assessment (KEA). The first is the School Readiness Assessment study—a one-time research undergoing that tracked the assessments of a sample of the kindergarten population—and an on-going K-2 formative assessment. The latter is akin to the K-3 assessment currently in production in that kindergarten, first grade and second grade teachers deliver it at the end of their respective years. Both have been used in conjunction for analysis of the public education system as is as well as to evaluate the needs of the school-age population. The goals of these two instruments were also to close the school readiness gap and improve student outcomes in early elementary grades. Although they do have significant weaknesses as assessment tools, they will be used by the think tank as a rough template to expand into the new K-3 assessment.

The inception of this movement to track the readiness of kindergarten age students was a recommendation by a 1999 School Readiness Goal team to assess incoming kindergarteners to understand their skills and knowledge better so that the system can be reformed to serve them. North Carolina thus gathered data for the School Readiness Assessment Study from 1,034 kindergarteners. The measures met all five domains of readiness and were appropriate for Children with High Needs. The findings indicated that on average, North Carolinian kindergarteners scored on par with the rest of the nation. However, lower income families scored significantly less than their peers from higher-income families. This study brought to light serious discrepancies in performances, and this achievement gap called for better transitions for children entering kindergarten as well as system reforms to address the inequality in performance among socioeconomic statuses (Scott-Little & Maxwell 2000).

However, this assessment study did have its limitations. It was a representative sample but not large enough to lend itself to be analyzed by various geographic district or counties. Further, it did not provide adequately detailed information on individual children to help guide classroom instruction. Although it is helpful to be able to undertake systematic reforms and larger policy changes, information can be used to target specific schools and classrooms to be able to better assist kindergarten teachers in their jobs. Finally, the data from this assessment and the formative KEA to follow did not contribute to a statewide longitudinal data system, which proved to have serious limitations in tracking trends over time and analyzing cause and effect scenarios to determine the efficacy of proposed reforms.

The KEA presented its own limitations as an assessment tool. In place since 1997, it was formed in response to the passing of legislation mandating individuated assessment instruments for first and second grades in lieu of the newly introduced standardized tests. While this instrument responds to the drawback of the School Readiness Assessment Study's inability to account for details of individual students, its targeted areas of assessment were limited to literacy and mathematic. The assessment was able to guide classroom instruction, but the state did not provide a system of uniform support and so teachers sought support from outside the realms of DPI, leading to inconsistencies across local districts. These inequities were exacerbated by the fact that the instruments to administer this assessment were left up to districts to decide in the first place, demeaning the value of a standardized assessment tool. Further, the state had not received the professional training to train teachers to use the results of the assessment to make instructional decisions. While the knowledge was available to inform reactionary changes, the tools were not ubiquitously distributed across the state, which limits the KEA's ability to close the achievement gap in a fair, egalitarian way (NCSL 2010).

National review of state approaches to school readiness assessment

School readiness has long permeated discussion among education providers and policy makers. Most stakeholders agree that readiness is more than a matter for schools but rather an issue that incorporates the joint effects of families, environments, and communities (NASBE 1991). School readiness assessments have been used to analyze risk factors and predictors of success in kindergarten-age students by screening for development in the five domains of: physical well-being and motor development, social and emotional development, approaches toward learning, language development and cognition and general knowledge (Kagan 1995). However, school readiness assessments vary in definition and application, and different states use them for different purposes.

These assessments have historically been used for five general purposes. The first type of assessment is one made to *improve learning*, and does so by gathering information on the strengths and weaknesses of students on an individuated basis to inform changes in classroom instruction and address particular needs. The second type of assessment has been used to *identify children with special needs*, and has generally been carried out in a two-step process. One step screens all children on developmental trends and cornerstones, and then the second step refers children with atypical or concerning behavior for a more thorough assessment specific to their needs in order to determine eligibility for special education services. The third type of school readiness assessment has been used to *evaluate programs*, more specifically, assess the positive impact of pre-kindergarten early childhood programs in terms of preparing children for the transition to school and minimizing the achievement gap upon entry to kindergarten. The fourth type of school readiness assessment has been used to *monitor trends over time* by providing a snapshot of children as they enter kindergarten and then again over the next several years of schooling to screen for trends. Unlike *program evaluations*, this method of readiness assessing focuses on the holistic impact of the public education system on the child rather than analyzing the impact of specific institutions on learning. Finally, the last type of assessment has been *use for high-stakes accountability*, which is a highly technical, rigorous mode of assessment to make concrete decisions about specific children or teachers. This type of assessment is considered high-stakes for the impact of results can lead to the letting go of a teacher or the prevention of a child from entering kindergarten even though they are of age (Scott-Little & Maxwell 2000).

Depending on the purpose of the assessment, school readiness assessments can track for several specific characteristics in kindergarten students upon entry that are strongly correlated to academic success and general readiness for school. The most comprehensive study that aimed to establish this set of characteristics was the Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K), which was a national study of a representative sample of 22,000 kindergarteners conducted by the U.S. Department of Education (Zill and West 2001). The study found that children from families that receive food stamps or temporary assistance, live in a single parent home, have a mother with less than a high school education or have parents whose primary language is not English are considered at-risk and generally score lower than average on the five assessment domains than their peers. The national study also found that girls, upon entering kindergarten, scored higher on reading skills than boys and about the same on math and general knowledge. The female gender was also correlated with better prosocial skills and a lower tendency to engage in problem behavior. Lee and Burkam (2002) used the national data of ECLS-K to ascertain that African Americans, Hispanic and other ethnic groups such as Native Americans scored lower on the five domains as populations than their White or Asian peers.

Questions have been raised regarding the predictive validity of kindergarten readiness assessments. One study examined the use of kindergarten assessments to predict academic achievement through the third grade among 463 students from a mid-sized Florida school district. The study used the Early Screening Inventory-Kindergarten (ESI-K), a commonly used form of readiness assessment and Dynamic Indicators of Basic Early Literacy Skills (DIBELS) to gauge the readiness of kindergarten students and then used the

Stanford Achievement Test-10 (SAT-10) and the Florida Comprehensive Assessment Test: Norm Referenced Test (FCAT: NRT) to evaluate academic achievement over the next three years of the tested kindergarten students. The study found predictive validity of the kindergarten assessment in overall academic achievement of students from grade 1-3, but found only moderate correlation with success on the SAT-10 and FCAT: NRT exams (Gillard, 2008). The study, however, was limited by its largely white and middle-class sample size. A broader meta-analysis of 70 longitudinal studies concluded that school readiness assessments only predict about 24% of the variability in children's kindergarten or Grade 1 academic and cognitive competency and about 7% of the variability in their social/behavioral competency (La Paro & Pianta, 2000). These results allude to the productive potential of kindergarten assessments but indicate that further development of the instruments is required to ensure greater predictive validity and functional use of results.

Teacher perceptions of kindergarten readiness can be used to augment the formal readiness assessments already being used in many schools as they have a more fluid understanding of children and the way they vary intra-classes and among various years. The Carnegie Foundation asked 7,000 teachers in 1990 about the numbers of students who were "ready" for kindergarten based on questions of students' cognitive, social, emotional and physical development. Overall, 35% of students were not ready to participate in kindergarten according to teacher perceptions of readiness, ranging from estimates of 23% in North Carolina to 47% in Hawaii (Boyer 1991). The diagnosis of readiness in children, according to kindergarten teachers, is varied but most teachers agreed on three key factors: 1.) Physical health, rest and nourishment, 2.) Communication of thoughts and needs in words, 3.) Curiosity and enthusiasm in approaching new activities. Recent data from the ECLS-K placed greater emphasis on nonacademic readiness skills than traditional domains of assessment. Specific tasks such as pencil holding, color recognition and shape identification were rated essential or important by less than one-third of teachers surveyed whereas over 75% of the 3,305 respondents ranked being able to follow directions and communicate needs as more important in terms of school readiness (Lin, Lawrence and Gorrell 2003). However, when looking specifically at the perceptions of kindergarten readiness among teachers in urban areas of low-income students, academic skills were ranked higher in terms of importance for school success than social skills (Wright, Diener and Kay 2000). Although principals in these areas continued to believe social skills to be paramount, teachers emphasized literacy for their students at risk for school failure. The impetus for policymakers is to keep in mind that kindergarten teacher perceptions of readiness may vary from the formal tools currently being used, and they may also differ among themselves depending on demographics of their students.

Currently, 25 states in the U.S. use kindergarten readiness screening and about half of them go beyond general cognition and literacy to address elements of the five domains of readiness. Of the 25 using readiness assessments, 21 are universal and 3 more are considered universal in practice, meaning they do not target select student populations or programs. Universal testing allows for the state and local community to take ownership of the results but also provides the opportunity for individuated alterations or instructions for schools and teachers to address specific needs at the child level. There are 20 states that require a particular assessment instrument, or limit the possible instruments to a short-list, and of these, 11 instruments evaluate between five and nine domains of readiness as opposed to the remaining nine instruments that simply evaluate reading. Of the 11, six use DIBELS, a common assessment tool that is publically available, focusing primarily on basic cognition and literacy. North Carolina is one of the states that utilize a state mandated assessment instrument but only tests for reading and math skills (NCSL 2010).

Many states use a basic teacher developmental skills checklist that is standardized and norm referenced based on observations of the child. However, it has been found to be vulnerable to teacher biases of gender, race and ethnicity, skewing the results and demeaning the validity of the assessment. Another mode of assessment is called the Teaching Strategies GOLD instrument and is based on observation and documentation of students on a naturalistic scale as opposed to standardized assessments. This is another commercially available assessment tool based on a 1-9 scoring rubric of readiness items, but is again based on observation and can be suspect to teacher biases. Finally, the Work Sampling System (WSS) is another naturalistic assessment instrument based on observations and a checklist, scoring the child as proficient, in process or not yet in the domains of kindergarten readiness. These commercially available tools have been used in many states but suffer from a dearth of empirical evidence of predictive validity. Yet, they are applied ubiquitously across the state allowing for uniformity and the comparison of districts against each other by the state. About 20 states indicate some level of reporting to the state yet others states that mandate the assessment of kindergarten students simply leave the method up to the local districts and do not require that the results be reported back to the state, precluding the potential intra-state comparison of districts. Of the 25 states requiring assessment, 12 do not publish any results and of the 13 that do publish their findings, only four publish state-level data (NCSL 2010).

It has been found that state-produced assessment instruments tend to have better predictive value as they can account for state specific demographic characteristics and learning standards. Maryland implements a hybrid model, fusing the Work Sampling System with state created domains of assessment. The Maryland Model for School Readiness (MMSR) assesses for reading, math, science, social studies, arts, physical education/health and personal and social development. The standard is meant to be a measurable statement of what students should know and be able to do upon entrance to kindergarten. The state used WSS as a basis but created their own set of content standards, which go beyond WSS to include historical, cultural and social context as well as creative expression and production. Teachers administer the assessment in the first quarter on all kindergarten students, electronically but the process is also continued with the creation of a portfolio assessment of work samples and observed documentation of learning, as recommended by WSS, with a "proficient," "in process," or "needs development" scoring system. Teachers must use this ongoing portfolio in conjunction with the assessment to monitor progress and implement instructional changes, however, only the first quarter reports must be submitted to the Maryland State Department of Education. The MSDE will create a data file on all students with and carry out their own analyses according to the MSSR construct, continuing along the way to perform reliability tests. Groups of students will be characterized as full, approaching, or developing in terms of school readiness. The ultimate reports produced with these analyses are disseminated to Local Education Authorities (LEA), policymakers and the early childhood community in the form of a cumulative annual report.

Methods

This phase of the project was meant to survey kindergarten teachers on the information about students they thought was most important as knowledge that would facilitate their ability to meet children's individual needs and to predict the students' academic success. Teacher input is necessary for two key reasons. First, they are uniquely engaged in the development of children, interacting with them in a classroom environment. This position allows them to offer insightful feedback on ways to improve the instrument and their approaches to the process of assessing kindergarten-age children. These opinions and perspectives offer significant value in considering the features the final product. It is possible that this population expresses different priorities than the academic literature on assessment validity, kindergarten readiness, and classroom instruction would suggest. Second, kindergarten teachers are going to be directly involved in the process of delivering and responding to these assessments. Their professional support is absolutely crucial to the effective and appropriate implementation of the assessment.

This goal was reached through a cross-sectional survey study that gathered teachers' valuations of the importance of each item suggested for inclusion on the final assessment tool. The mean valuations of each item would be calculated and used to rank the suggested assessment items in order of importance according to teachers. This information was intended to help the think tank understand the opinions of kindergarten teachers on the content of the assessment as well as to allow the kindergarten teachers to provide feedback in terms of comfort level in administering the assessment in the home, possible negative implications and their personal suggestions.

Instrument

The survey included 22 items suggested for inclusion on the K-3 assessment (see Appendix 1) to assess student performance in each of the five domains mandated for evaluation by state legislation. The survey asked the teachers to rank the items on level of importance using a Likert scale ranging from 1 (of least import) to 6 (greatest import). This scale was intended to encourage mean rating of items spanning a spectrum of importance and so respondents were asked to seriously consider using each of the 1-6 options to prevent clustering on either the upper or lower end of the scale. A 1-5 scale was not used to avoid the consistency of neutral answers of "3." Each of the items was to be valuated individually, not in reference to the previous or latter items; therefore, we did not limit how often the respondent could use a particular option in answering the questions. The survey also asked the teachers to evaluate their level of comfort, on the same 1-6 scale, of administering the assessment in the homes of children and of using an electronic tablet to do so. Finally, teachers were asked to indicate the likeliness that the suggested assessment items on the survey would equip them with knowledge that would hinder their ability to address their students' needs appropriately and fairly. The median score of 3.5 on this 1 to 6 scale was used as a benchmark against which to compare each mean response to the suggested items and assessment questions.

Participants

The population for the study was the totality of kindergarten teachers working in public schools in the state of North Carolina. At the start of 2013, within North Carolina, there were 9,493 kindergarten teachers working in public schools. The sample for this study was a non-randomized subset of the population that voluntarily completed the survey. The study was meant to allow each kindergarten teacher to have a voice in the process of creating the assessment tool, as they will all be affected by its implementation

Further, the survey was meant to be anonymous, with the only identifying questions being the district the respondent worked in, the number of years they worked as a kindergarten teacher and their highest degree obtained. Creating a representative random sample would have compromised the integrity of the survey as anonymous and possibly led to less than truthful answers by teachers uncomfortable with being identified by their responses.

The survey template was on an online data collection tool known as Qualtrics. The items were randomized to prevent framing and biases. This online survey was made available through a link and made private, prohibiting anyone without the link from viewing the survey, let alone taking completing it. The link was administered by email to the various teachers when the survey went live on February 20, 2013. For the sake of privacy, the email was not sent directly to the inboxes of various teachers. Instead, an email was forwarded to the principals of all public elementary schools in North Carolina, asking that they forward the email and attached survey to the LISTSERVs of kindergarten teachers at their schools (see Appendix 2). The duration of the survey was 20 days and it closed officially on March 2, 2013.

The final sample was analyzed in terms of, identifying characteristics such as distribution by size of locale, geographic location, and highest degree earned. The Department of Public Instruction has a database, which records these features for every kindergarten teacher working in North Carolina. The distribution by locale size, region, and most recent degree earned in our sample was compared to the distribution of these characteristics for all kindergarten teachers in the state to assess how representative our sample, roughly. The sample would be considered roughly representative of the state if the proportions of the sample's composition approximately match the profile compositions of all the kindergarten teachers in the state. Additionally, the sample was analyzed by distribution of number of years teaching. Although the state does not have this data to compare with our sample, we believed it was important to understand the experience of the teachers responding to the survey. It is possible that new teachers may value different items or feel differently about the overall assessment than veteran teachers.

Geographically, the regions were separated into Northeast, Southeast, North Central, Sandhills/South Central, Piedmont-Triad/Central, Southwest, Northwest and Western. City Size was separated into three groups. A final group was made for cases in our sample with unspecified regions, including charter schools and non-responses (Appendix 2). Respondents were to provide which district they work within and this information was used to separate the sample into subgroups by region as well as locale size.

To analyze distribution by locale size, we created four groups, which we fit various community sizes and types into, to be able to compare the effects locale size have on the responses. The state's data on all kindergarten teachers separated them into large cities, midsize cities, small cities, large suburban areas, midsize suburban areas, distant towns, fringe towns, and remote towns. For our study, Group 1 included all large cities. Group 2 included midsize cities, small cities, large suburban areas, midsize suburban areas, distant towns, fringe towns and remote towns. Group 3 included distant rural areas, fringe rural areas, and remote rural areas. A separate Group 4 was created for schools with unspecified locale sizes, including charter schools and non-responses. Charter schools were not surveyed because this assessment is to be administered in public schools (Appendix 2). The state's placement of various districts into groups of locale sizes was used as the basis of differentiating the sample cases into locale size based on their reported districts.

Due to the fact that teachers from various educational backgrounds or with different experiences teaching could have different opinions, we compared the distribution by degree and years teaching in our sample to the population. Teacher education degrees were separated into Bachelors, Masters, Professional, and Doctorate. Surveyors were asked to select their number of years teaching from 1-5 years, 6-10 years, 11-20 years and more than 20 years as seen in the survey itself (Appendix 2).

These measures were taken to ensure that no unique subset of the population had a disproportionate voice in our surveying. If a significant demographic difference between our sample and the population was demonstrated, the results would not be considered as truly representative of all the kindergarten teachers in the state. Discretion was used in the treatment of the data and measures were taken to assure utmost anonymity of responses.

Data Analytic Plan

The first statistical test completed was a basic analysis of the mean responses to each assessment item suggested for our sample. These responses were then ranked from lowest to highest mean response to see an overview of what kindergarten teachers deemed important to assess at the beginning of the school year and which items were not as essential to evaluate. The survey items were divided into the five domains recommended by the state and the think tank (see Appendix 3) to ensure that each of the domains was addressed, as the requirement for their inclusion must be fulfilled in creating this assessment. The domains were used to create super variables, or larger categories used by combining the responses of each of the items that fall within each domain and thus finding the mean rate of importance (on a 1-6 scale) of this conglomerate of ratings comprising the domains. These domains would then be ranked from highest mean value of importance to lowest.

The mean responses were calculated for level of comfort administering the assessment in a home to show the amount of training a teacher might need before administering the assessment effectively in the home. Further, the mean level of comfort of kindergarten teachers with using tablets to administer the assessment was evaluated. Given the efficiency and ease of tablets for collecting survey data, it is a viable possibility for the mode of implementation of the K-3 assessment, but it is imperative that a teacher feels comfortable with the technology before agreeing to this method of assessment completion. The means were also calculated for the amount teachers believed the information gained from the assessment could be helpful to them and whether or not the assessment would be more helpful if it included follow-up assessments throughout the year. Given that the assessment is meant to assist teachers in their instruction decisions, it is important to ascertain whether they would value and be able to utilize the knowledge the assessment results would give them.

Finally, mean responses were calculated for the likelihood that the knowledge the assessment results would give teachers would be harmful in educating their students. Studies have shown that observation-based assessments given by teachers have been biased by race, gender and ethnicity, among other factors and so, it is possible some of the sensitive knowledge obtained from this assessment could bias a teacher for or against a particular child. For this reason, we asked that teachers evaluate the likelihood of any information they would gain if the suggested items were included on the assessment hindering their ability to appropriately meet the needs of children in the classroom.

These general questions regarding the assessment and its distribution were analyzed further based on frequency of response choices from 1 to 6. The mean responses for the total sample and individual groups within the sample sufficed in illustrating the importance kindergarten teachers assigned to assessing the items suggested for inclusion on the assessment. However, for the creation and implementation of the assessment, it is imperative that teachers show support and willingness to carry out the state's plan. Therefore, we also looked at the breakdown of the frequency of each response for questions 24-28 (Appendix 1). For example, even if most teachers felt comfortable using a tablet, if a relatively significant minority rated their level of comfort at 1 (the least comfortable), then the think tank should consider offering training to teachers before fully implementing the assessment statewide.

Beyond these basic tests, various other statistical analyses were run to compare the responses among the different subsets of our sample. Using ANOVA, we were able to find the questions that different groups of teachers (by years teaching, region, degree, or locale size) had statistically significant variations in mean responses. Based off of these results, we looked further into the mean responses for each group for each of the questions they differed in to evaluate how robust the difference in means were. Any interesting finds were then analyzed to find trends in which groups disagreed on which items. For example, groups 1-3 for locale size were compared to each other using t-tests to ascertain which group differed from the rest on the questions that had significant ANOVA results. The teachers were asked to select the number of years they have been teaching between four groups (Appendix 1). These groups represent a continuous variable and so a Pearson correlation test was used to find any significant differences in opinions between the groups. The focus of these demographic analyses was on questions regarding the assessment instrument itself and its distribution. Any finds that were not found to be too significant were summarized in charts and included in the appendices.

The survey also contained two sections where respondents would be allowed to offer their suggestions or opinions regarding the composition of the assessment or its administration. Given the varied nature of comments, we would deal with them by coding them into groups with similar themes to be able to roughly quantify the major concerns or suggestions teachers had. This process would only identify and group unique responses that offered ideas or reservations that were not already expressed in the evaluation of assessment items or the second part of the survey, involving questions about the instrument itself. For example, a comment explaining a rating by introducing information not readily apparent in the means calculated for each of the questions. The remainder of the comments would be grouped into an “other” category, including those of general support or dissent on the survey.

Hypotheses

Given the existing literature in the field and the previous studies that have been done, we believe that the items in the Emotional/Social Development domain would be the highest ranked, followed by those in Approaches to Learning. This hypothesis is based on the recent emphasis teachers have put on intellectual curiosity and enthusiasm as well as the ability to communicate thoughts and needs. However, we also believe that if analyzed by separate populations, Group 1 (large cities) would place greater import on Cognitive Development than the other four domains, given studies that have shown that teachers in urban areas with students coming from low-income households continue to find cognitive abilities the greatest indicator of later academic success.

In terms of the actual assessment instrument, we hypothesized overall support for the tool and the proposed items. We foresaw some hesitation in conducting home visits, and believe that these hesitations would be more pronounced in teachers from large urban environments. We also believed the notion of utilizing a tablet would be looked upon favorably. We believe that teachers will not find the information gained from this assessment harmful to their ability to attend to their children's needs.

Results

Representativeness of Sample

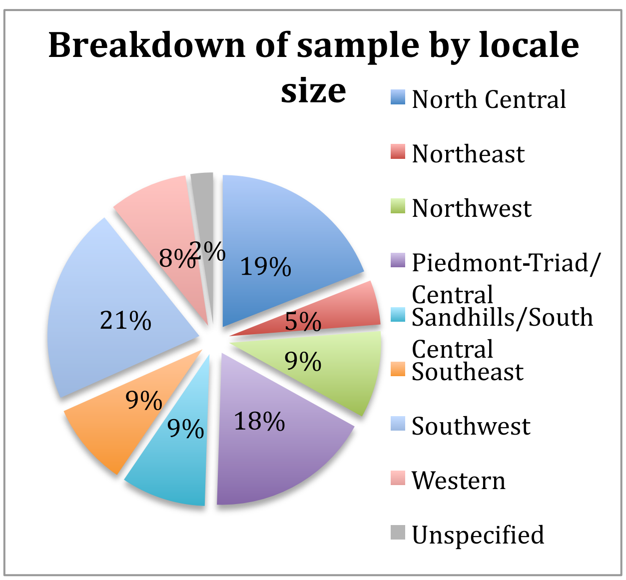
Our survey received 2,596 responses or 25.62% of the total number of public school kindergarten teachers in the state. Of these, there were some incomplete surveys that filled out some basic demographic information, such as district or number of years teaching, yet did not continue to finish the whole survey. We found 151 survey responses that did supply demographic information about their district, experience and most recent degree, yet did not complete the rest of the survey. Another 22 cases did not include information about the school district they work within yet did answer questions about experience and most recent degree. Of these 22 surveys, however, 8 did still complete the remainder of the survey questions. The study had 2,432 complete surveys that filled out all of the demographic information and responded to each of the survey questions. A complete list of the schools the respondents worked for can be found in the Appendix (Table 1).

Each of the tables outlining the responses and test results also include the number of responses included in the evaluation process. Given the large sample size, a handful of missing responses can be considered a minor error and does not skew the validity of the results for our purposes. The means, in spite of the discrepancies in numbers of cases evaluated, still stand to give a general idea of the opinions kindergarten teachers in North Carolina have about the assessment items, the instrument itself and the information it will be used to gain.

For the remainder of the paper, “sample” will be referring to the 2,596 responses to the survey and “population” refers to the 9,493 kindergarten teachers working in public schools in North Carolina. First and foremost, our sample was assessed to determine if it could be considered a general representation of the population. While our sample is not probability sample and was not randomly selected from the population, it does have certain qualities that lends to being comparable to the total population.

The sample response was quite similar in distribution by region to the actual population of kindergarten teachers in North Carolina (Fig 1.) For each of the regions, the representation in the sample never varied by more than two percentage points from the representation of that region in the population. For instance, 21% of our respondents stated that they were from schools in the Southwest region, and the Southwest region has

21% of all Kindergarten teachers in North Carolina. The Piedmont-Triad/Central region was the same percentage portion of our sample as they are of the total population. The Sandhills/South Central, Southeast, North Central, and Northeast region groups within our sample were within 1% of the percentage of the population they represent and the Western and Northwest groups were within 2% of the percentage of the population they represent. Our sample also accounts for a 2% error group of respondents that were either from charter schools or chose not to respond with which district they work within. The difference between the makeup of the sample and population by region was not statistically quantified but seems not to be substantial in terms of the purpose of this study. No region seemed to have a disproportionate contribution to this survey of the opinions of kindergarten teachers.



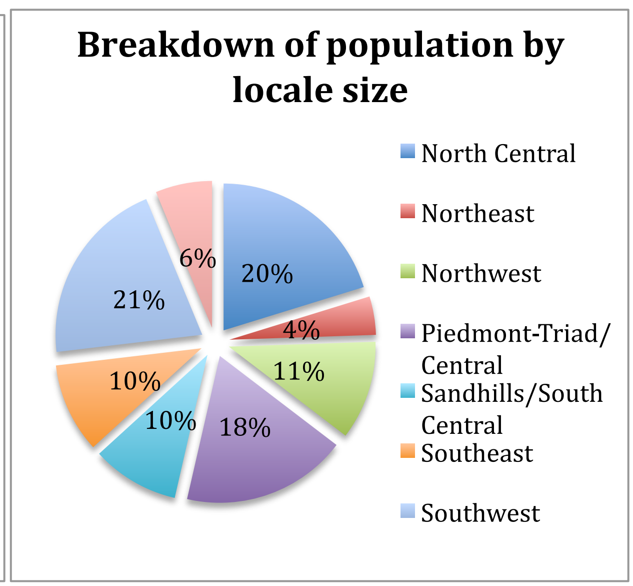


Figure 1: Comparing the distribution of the sample to the population based on regions

The sample was also comparable to the population in terms of distribution of locale size (Fig. 2). The locale sizes were broken down into four groups as explained in the methods section. Our sample comprised of 22% respondents from Group 1, 32% respondents from Group 2, 44% respondents from Group 3 and 2% respondents from Group 4. Groups 1 and 3 matched the population exactly in terms of percentage, and Group 2 differed by 2% from the proportion Group 2 represented of the total population, based on state data. Our sample included an error of about 2% due to cases of unspecified school or respondents from charter schools. These numbers suggest that the possibility of locale size being a confounding variable in our results is minimal due to the fact that each the proportion of each group in our sample seemed to roughly mirror the proportion each group contributed to the total population.

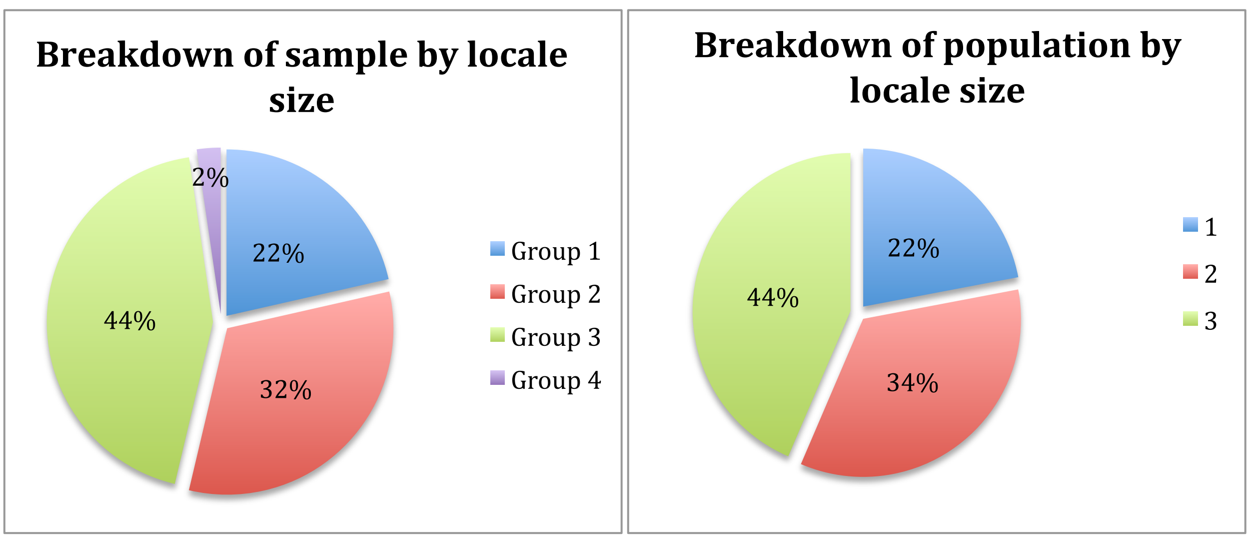


Figure 2: Comparing the distribution of the sample to the population based on locale size

The sample did not demonstrate a large variation in the distribution of education level within the sample, compared to the population (Fig. 3). In North Carolina, 73% of teachers have a Bachelors degree and 27% of teachers have a Masters degree. There is a negligible amount of teachers who have Advanced or Doctorate degrees. In our sample, 67% had a Bachelors degree and 33% had a Masters degree. There were negligible cases within the sample of respondents with Doctorate or Professional degrees. While these numbers do not exactly match the distribution of degrees in the population, they are approximately similar and so we treat the sample as relatively representative of the population in terms of education.

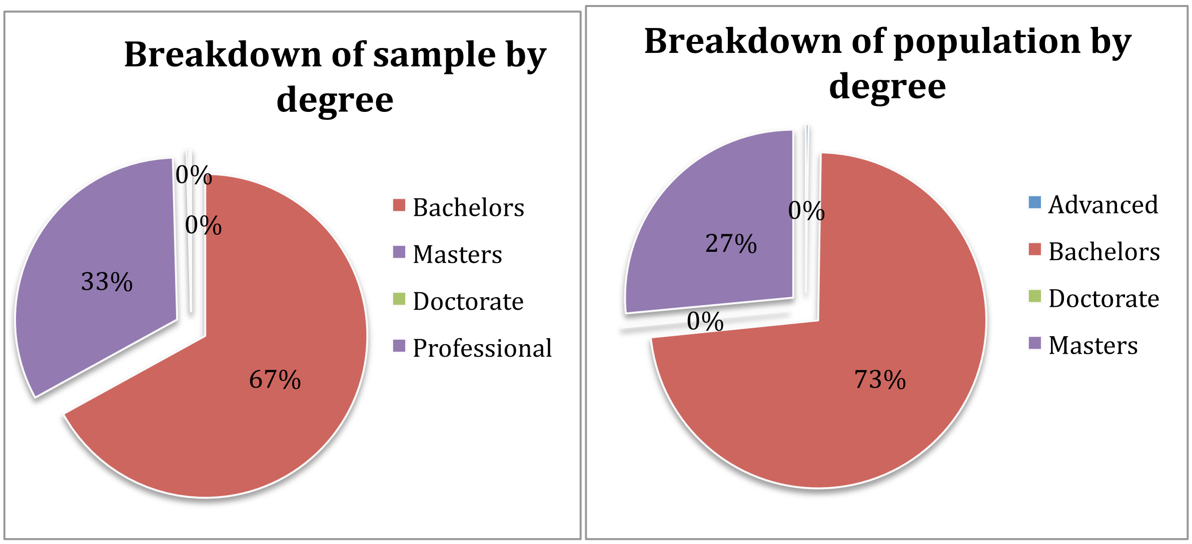


Figure 3: Comparing the distribution of the sample to the population by highest degree earned

The state does not have information regarding how long each particular kindergarten teacher within public schools has been teaching (Fig. 4). However, we still analyzed the distribution of the sample in terms of experience to ensure that the respondents were not entirely comprised of new teachers or veteran teachers. Difference in experience can cause difference in opinion thus an ideal sample would include the voices of teachers with various levels of experience. Of our sample, 27% of respondents had been teaching from 1-5 years, 22% from 6-10 years, 31% from 11-20 years and 20% for over twenty years. We do not know the proportions of teachers by years of experience for the population, but this data about the sample shows that no group was overwhelmed by the opinion of another or went entirely unheard in this survey.

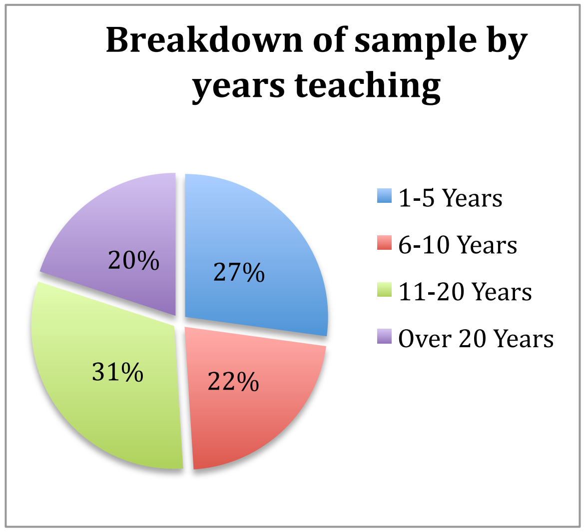


Figure 4: Breakdown of sample by years teaching

Teacher Ratings of Domains

The responses to the survey were first evaluated as a whole sample to find the aggregate opinions of all the kindergarten teachers in North Carolina. The survey consisted of several items suggested for inclusion on a K-3 assessment, and the teachers were asked to evaluate the importance of validly testing each of these items for their day-to-day work of instructing a class. The mean rating of importance for each of the items was calculated, and the items were then ranked in ascending order of importance (Fig. 5). For these calculations, we included the means of the 2,432 cases that had completed this part of the survey (Table 2.1). On a scale of 1-6, with 1 being of least importance and 6 being of most importance, a mean rating of 3.5 would be considered an average rating of neutral importance. The individual responses were varied, ranging from 1 to 6 for each of the items (Table 2.1). The respondents as a whole found each of the 22 items to be of above average importance, with the exception of Parent's Education Level, indicating that nearly all the items suggested for inclusion on the assessment are considered important, though at varying degrees, for a teacher to know in order to meet their students' needs.

Of the top five highest rated items, three involved language. Receptive language received the highest mean rating of 5.36, followed by Expressive language and Developmental concerns, both with a mean rating of 5.23. Literacy followed with a mean rating of 5.18 and then Ability to comply with adult requests had a mean rating of 5.12. Parent's education level received the lowest rating, with a mean of 3.08. Family health history was the only other item to receive a mean rate of less than 4 also, with a mean of 3.57. This pattern suggests that kindergarten teachers do not put much emphasis on the knowledge of a child's background yet consider a child's literacy and communicative abilities to be integral to academic success.

While language seemed to dominate the top-rated items, these results also bring to light the validity of all five domains of assessment as generalized areas of development (Appendix 3). Items that fall within the same domain did not seem to cluster in terms of all having similar ratings of importance. Health and Physical Development as a domain, in particular, seemed to have particularly varied ratings. Fine motor skills was ranked 9th in terms of importance, yet gross motors skills was ranked 14th. Further, family stability was ranked 16th, yet family health history received one of the lowest rankings of 21st (Fig.5). This pattern could indicate that the manner in which this study separated items into domains did not accurately represent the way a kindergarten teacher would group the items. Or, this could suggest that at a kindergarten age, teachers do not expect that every facet within a particular domain of development needs to be assessed.

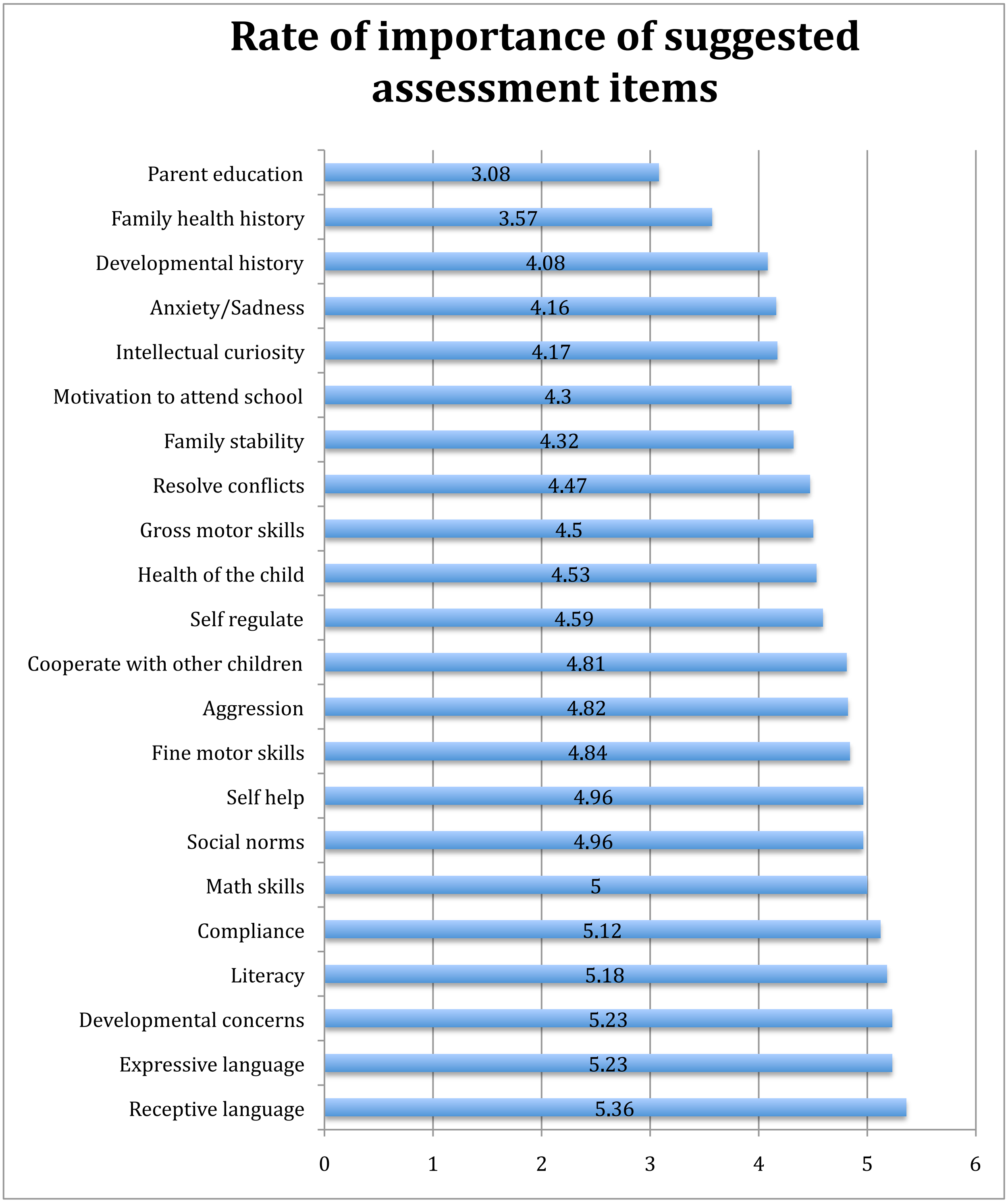


Figure 5: Ascending rank of importance of suggested assessment items

Each of the items was coded into one of the five domains of assessment prioritized by the state (Appendix 3). For each case, a mean was calculated by averaging the ratings of items within each domain, for all five domains. There were cases with an average rating of 1-6 for each of the domains (Table 2.2). The sample mean rate of importance was found for each of the domains by averaging the individual case ratings of domains, and then they were ranked. Language Development and Communication was ranked highest, with a mean of 5.26, followed by Cognitive Development with a mean of 5.15, Emotional-Social Development with a mean of 4.7, Health and Physical Development with a mean of 4.31 and finally, Approaches to learning with a mean of 3.85. Each of the domains received a mean rating higher than 3.5, and so it seems kindergarten teachers deem all of them important to assess in incoming kindergarteners (Figure 6).

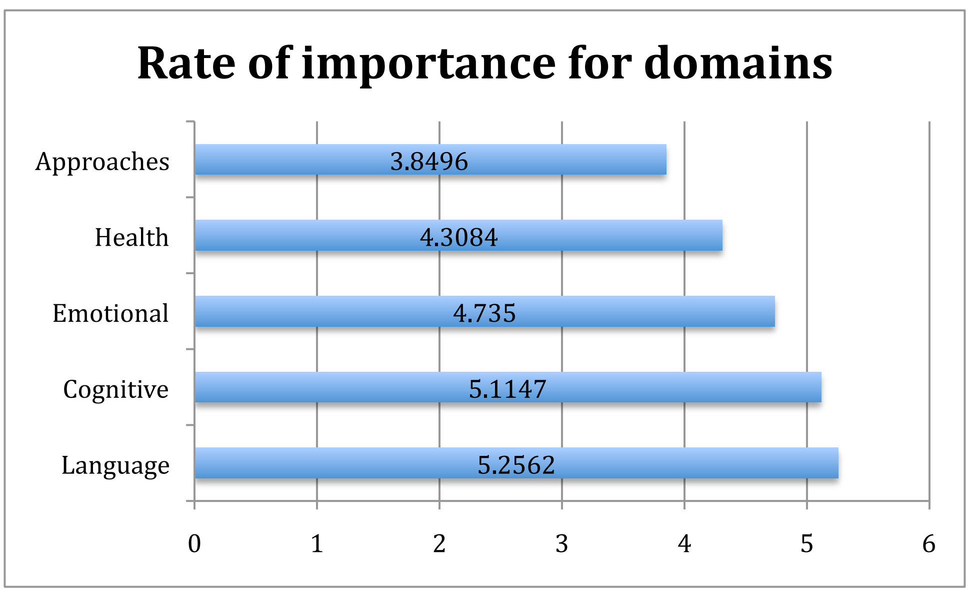


Figure 6: Ascending rank of importance of domains of assessment

The second part of the survey included five questions regarding the instrument and the administration of the assessment (Appendix 3). This section was analyzed using all available data, and so the number of cases averaged varies between the five questions, depending on how many answered each. We found a spread of responses to each question ranging from 1 to 6 (Table 2.3).

Tablet use. For the sample, the mean rate of support for using a tablet was 5.17 on a 1 to 6 scale, 1 being least support and 6 being most support (Figure 8). The respondents expressed overall support for using a tablet to administer the assessment, above a mean rate of 3.5, which indicates neutrality. Looked at further, it is apparent that 57%, over half the sample, expressed utmost support for the use of a tablet. However, 3% selected option 1, expressing the least support and 3% selected option 2, and 6% selected option 3. Options 1 to 3 show varying degrees of lack of support, which shows that approximately 12% of the kindergarten teachers in North Carolina do not feel comfortable with using a tablet to administer this assessment (Fig. 7).

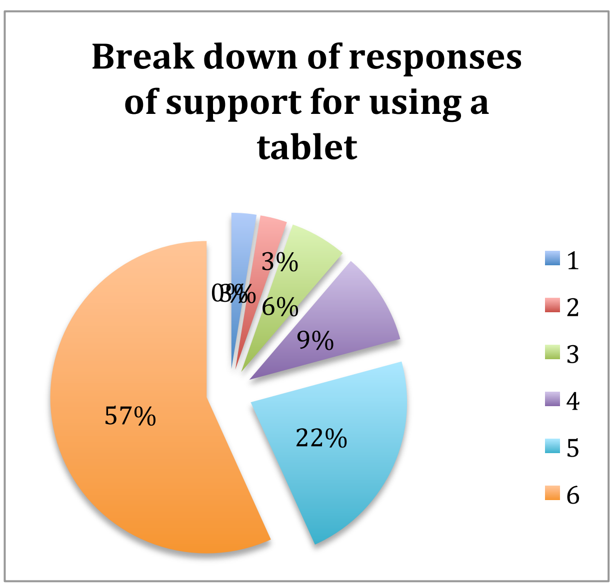


Figure 7: Distribution of responses for comfort using a tablet

Helpfulness of assessments. The teacher's rated the helpfulness of assessment information as 5.06 on a scale of 1 to 6, 1 being least helpful and 6 being most helpful. There was relatively less support of the notion of conducting multiple assessments in the year. However, on the whole, there was still general support for the suggestion that multiple assessments each year could be helpful to kindergarten teachers in their function. The mean evaluation of how helpful conducting multiple assessments would be was 4.29. Finally, teachers rated the likeliness that information gathered from this assessment would be harmful in the instruction of children at 2.35 on a 1 to 6 scale, 1 being least likely and 6 being most likely (Fig. 8). This was the lowest rating of any of the questions regarding the instrument or the delivery of the assessment and indicates that kindergarten teachers do not deem it very likely that the assessment could have negative side effects.

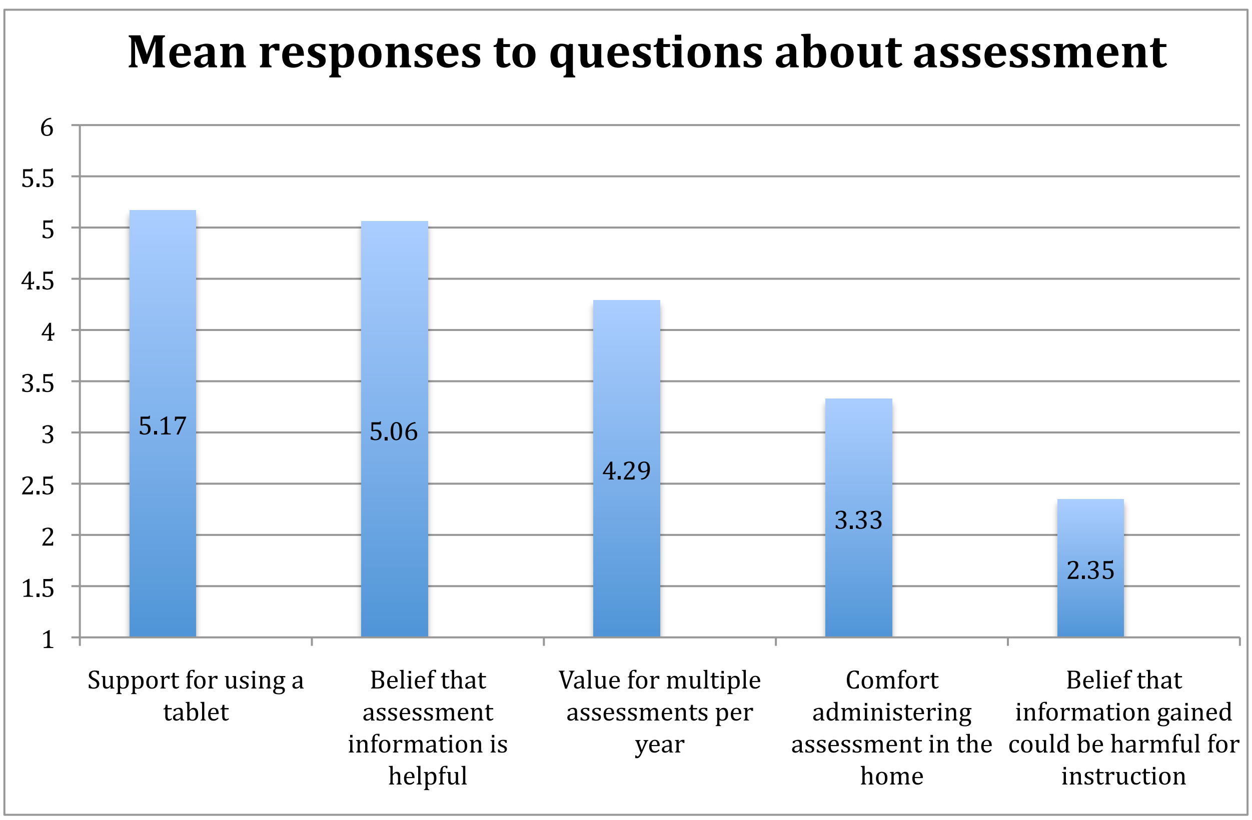


Figure 8: Mean responses to questions about assessment instrument and delivery

On the whole, teachers did believe that conducting multiple assessments in the year could be helpful in meeting the needs of a class and informing their instruction, with a mean rating over 3.5 indicating neutrality. More than a quarter of the sample, or 26% rated the helpfulness of multiple assessments at 6, the highest option. Another 26% rated helpfulness at 5, and 20% rated the helpfulness at 4, indicating that 72% of the sample express some level of support for conducting multiple assessments in the year. However, 7% rated multiple assessments as not helpful at all and 8% deemed multiple assessments as barely helpful. A full 28% do not support conducting multiple assessments per year (Fig. 9).

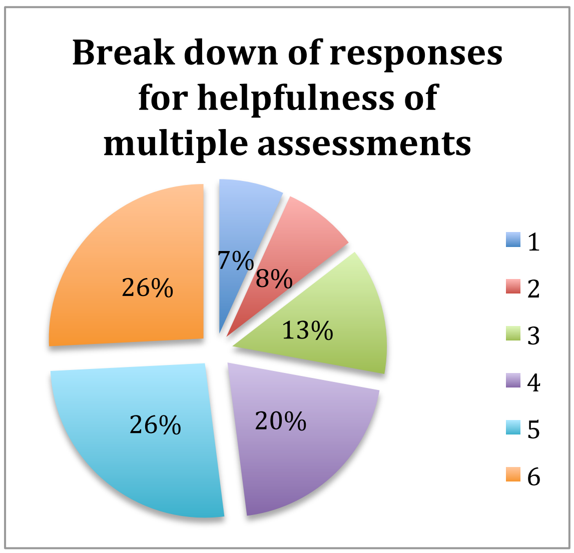


Figure 9: Distribution of responses for helpfulness of multiple assessments

The mean rate of likelihood that the assessment information is harmful was far below the neutral rate of 3.5 on a 1 to 6 scale, 1 being least likely and 6 being most likely. Of the sample, 32% believed it was not at all likely that the assessment information would be harmful and chose option 1. Another 30% chose option 2 and 19% chose option 3, which shows they also believed it was not likely that the information would be harmful. Cumulatively, a full 81% do not believe it is likely that the assessment would have negative implications for classroom instruction.

However, 19% still harbor belief that the assessment could be harmful in meeting the needs of children in the classroom. Of this group, 3% selected the option that it was most likely that the assessment information could be harmful and 5% chose option 2, believing it was very likely that the information could be harmful (Fig. 10).

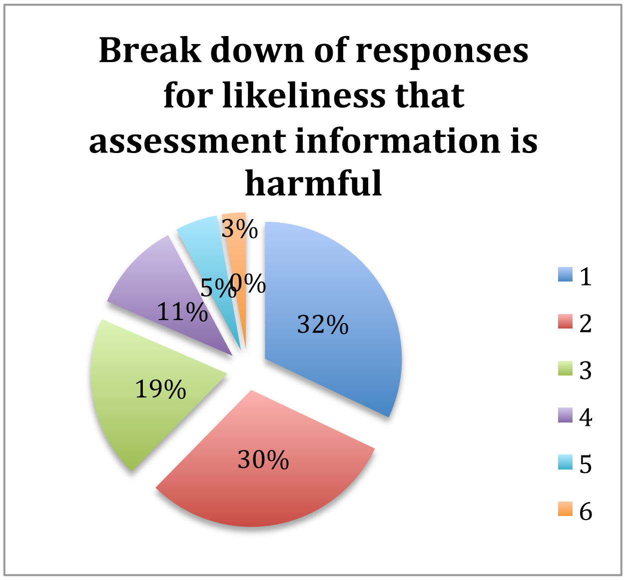


Figure 10: Distribution of responses for likeliness that information is harmful

Home visits. The least support from the sample was expressed for the notion of administering the assessment during home visits. On a scale of 1 to 6, one being least comfortable and 6 being most comfortable, the mean rating was 3.33, indicating a less than neutral rate of comfort level. Broken down, 24%, almost a quarter, of the respondents chose 1, which expressed the least level of comfort with conducting home visits. Additionally, 14% chose option 2 and 15% chose option three, which shows that a full 53%, or over half of the sample, do not feel comfortable with conducting home visits (Fig. 11).

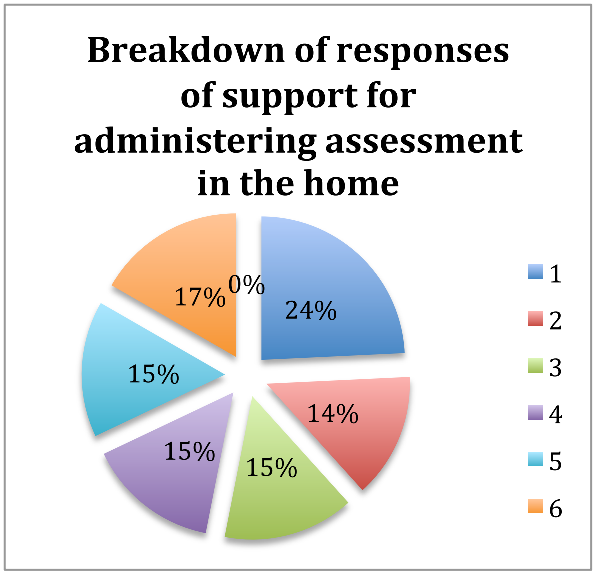


Figure 11: Distribution of responses for willingness to conduct home visits

Subgroup Differences

Statistical analyses exposed significant differences in mean rating of questions regarding the assessment and its delivery among various subgroups of our sample. One-way ANOVA showed variation among locale sizes regarding comfort conducting home visits (Table 4.1). For Group 1, the mean rate of comfort was 3.18, for Group 2, the mean rate of comfort was 3.53 and for Group 3, the mean rate of comfort was 3.25 (Table 4.2 and 4.4). Group 4 was excluded for its non-specificity and small sample size. A basic student's T-test for equality of means showed that there was a significant difference between the mean rate of comfort of Group 1 and Group 2, with a mean difference of -.344, with a p-value of .001, indicating statistical significance at an α-level of .01 (Table 4.3). A second t-test for equality of means found that there was a significant difference between the mean rate of comfort of Group 2 and 3 in conducting home visits. The mean difference was .281, with a p-value of .001, indicating statistical significance at an α-level of .01 (Table 4.5). One-way ANOVA tests showed that no other questions regarding the assessment instrument and its delivery varied among the subgroups of locale size.

A one-way ANOVA test also showed a significant variation in means between groups of teachers by years teaching on their rating of the helpfulness of multiple assessments, comfort using tablets, and support for conducting home visits (Table 5.1). Moving from Group 1 to Group 4 illustrated a steady decline in mean rating of comfort using a tablet (Figure 12). Because the groups of years teaching represent a continuous variable, a Pearson correlation test was done to analyze the relationship between the experience of a teacher and their rating of the survey questions.

There was a statistically significant relationship between years teaching and comfort using tablet, with a correlation of -.095, and a p-value < .001(Table 5.3). Briefly, this shows that a movement from Group 1 (1-5 years teaching) to Group 2 (6-10 years), shows a .095 decrease in rate of comfort using a tablet to administer the assessment (on a 1 to 6 scale). Less experienced teachers are more comfortable with tablets.

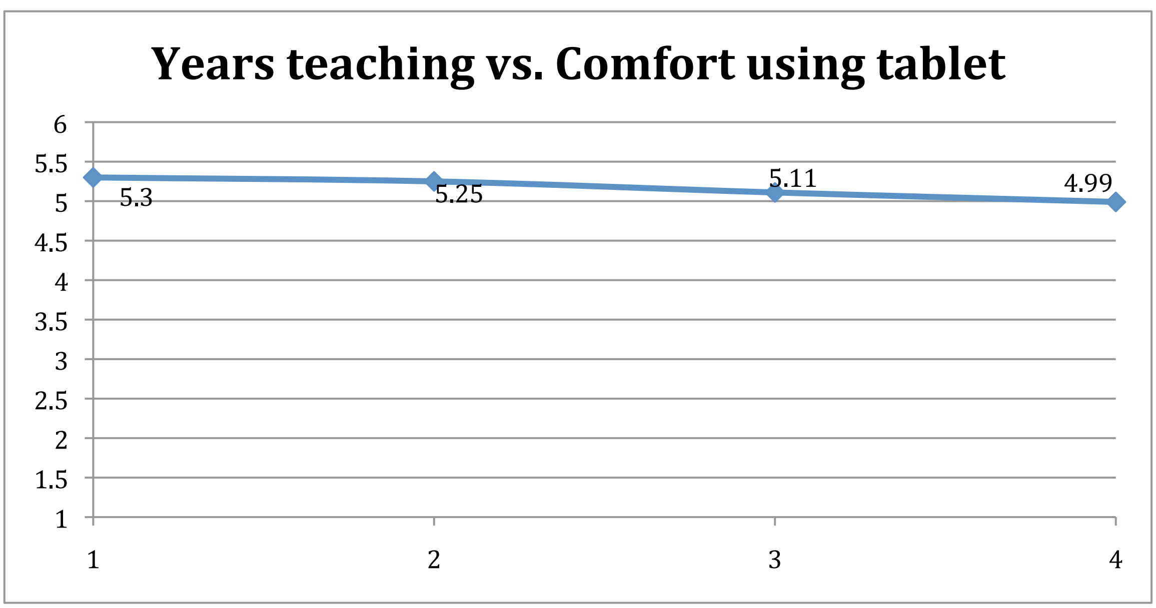


Figure 12: Graphing mean comfort using tablet against years teaching by group (1-4)

Another Pearson correlation test found a correlation of -.059 between years teaching and comfort conducting home visits, with a p-value of .004, considered statistically significant at an α-level of .01 (Table 5.4). This shows that as teachers gain more experience, or move to a higher group, they tend to be less comfortable with conducting home visits. However, this is not a consistent decline, as the mean rate of comfort is not steadily declining from Group 1 to Group 4 (Figure 13).

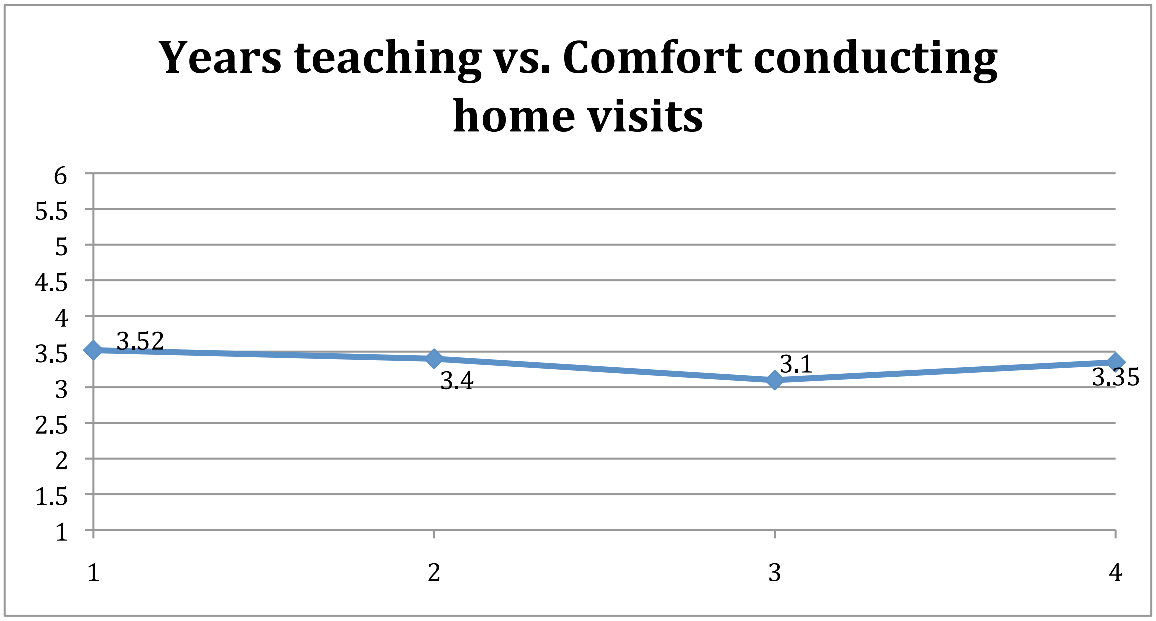


Figure 13: Comparing mean comfort conducting home visits against years teaching group (1-4)

Finally, a Pearson correlation test was also used determine whether there is a significant relation between years teaching and rating of helpfulness of conducting multiple assessments in the year. The correlation was found to be -.059 at a p-value of < .001 (Table 5.5). This relation outlines a decreased likelihood of finding multiple assessments helpful as teachers gain more experience teaching (Figure 14). However, this correlation and the comparison of mean ratings of helpfulness between these groups show that while the relationship might be significant, it is not very robust.

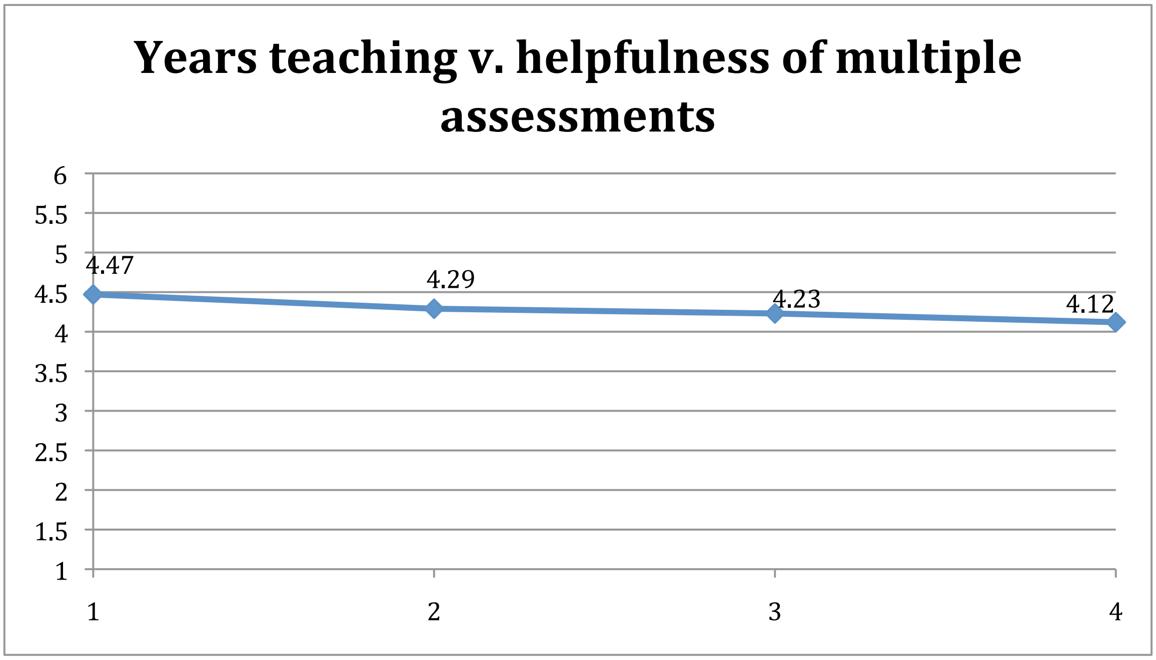


Figure 14: Graphing mean rating of helpfulness of multiple assessments against years teaching group (1-4)

A one-way ANOVA test also revealed significant variation among regions in their mean rating of comfort conducting home visits as well as their belief that conducting multiple assessments each year would be helpful (Table 6.1). The ANOVA test found no other significant variations among regions in their mean responses to questions from the second part of the survey, regarding the instrument and its delivery. Upon comparing the means among regions of multiple assessments and home visits, it is apparent that increases and decreases in means for multiple assessments among regions seemed roughly parallel to increases and decreases in comfort conducting home visits among regions (Table 6.2). Analyzed further using a Pearson correlation test, multiple assessments and home visits has a correlation of .385 with a p-value of essentially zero (Table 6.3). This shows that those more likely to feel comfortable conducting home visits would also find multiple assessments more helpful and vice versa.

Tables 7-13 include the significant findings from conducting ANOVA tests comparing the ranking of items and domains by years teaching, degree earned, locale size, and region. Only the items and domains with significant variations (at an α-level of .01) across their respective subgroups were reported. These findings show that the individual characteristics associated with urban environments or many years of experience teaching can lead to variations in opinions regarding the importance of assessing various areas of development. However, none of these variations were too pronounced or robust enough to merit much discussion.

Comments

The survey also included two areas for respondents to be able to include their opinions, suggestions or comments to expand on their survey answers. Following the first 22 items surveyors were asked to rate for importance, there was an opportunity to offer comments. In this field, 864 respondents chose to comment (Table 14.1). Of these, several contained typing errors or simply stated that the respondent had no comment. Further, the comments that expressed general support for the assessment or any of the items were also placed in the category of "other." The "other" category comprised 36% of the first set of comments. Of the remaining comments, each new idea was put in a separate group and its proportion of the total comments was found by the number of times it was mentioned over the total comments made.

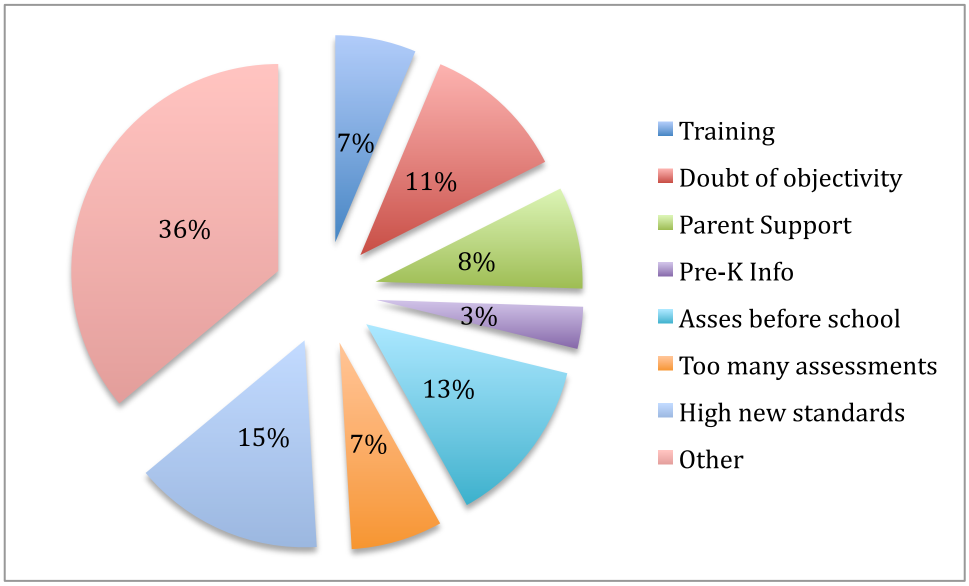


Figure 15: Breakdown of comments regarding assessment items

Several major concerns raised regarded the challenges presented by the new higher standards mandated by the Common Core, which accounted for 15% of all the comments (Fig. 15). These comments raised concerns about the emphasis on academics without placing equal import on development as well as the need to "hit the ground running" in order to cover all of the material necessitated by the curriculum. One surveyor noted "Children really need to be ready to enter kindergarten since the core curricula are so demanding." Another noted that the assessment tool should be used to determine whether or not a child is ready for the rigorous new standards of kindergarten: "If the children do not reach a certain level of qualification for kindergarten, they should be permitted or required to attend Pre-K for a year before entering kindergarten."

Another 13% of comments mentioned the importance of administering this assessment before the start of the school year (Fig. 15). These comments ranged from simply stating the importance of a teacher having access to this information about their class before meeting them, while others called for a staggered entry system to allow teachers time to conduct the assessments. One teacher stated, "I feel that children who have developmental delays need to be identified BEFORE coming to Kindergarten. There needs to be support for their educational and other needs in place before they enter kindergarten especially with the new curriculum."

From the sample, 11% of the comments expressed doubts about the objectivity of the assessment in its ability to assess a child's emotional or social development fairly (Fig. 15). They noted that children may behave differently in their first encounter with a teacher administering an assessment or that a child's true social ability would not be apparent until they are observed engaging with classmates. Further, some teachers brought to light concerns about the validity of forms filled out by parents. One teacher commented, "We have found that Social History reports, filled out by parents, are not often an accurate assessment of their child's abilities. It is important to know about traumatic events, possible pre-school concerns (behavioral/academic), and testing completed prior to kindergarten."

Despite these concerns, 8% of teachers in the sample still mentioned the importance of knowing a parent's level of support and willingness to partake in their child's educational experience (Fig. 15). Issues of participation included reading to children, speaking to them in English in a bilingual family, helping them with homework or attending school events. One surveyor mentioned, "Children come to us in all sorts of situations. It is important that parents be held accountable for their child when they start kindergarten and when there is academic concern."

About 7% of the sample mentioned the desire or need for training prior to administering this assessment (Fig. 15). This was a broad topic, but some of the comments discussed the lack of comfort teachers felt in delivering this assessment. Some mentioned they would need training before being asked to evaluate children on the suggested items or they preferred that trained social workers or other professionals in that field administer the assessment. Still others expressed the need to educate teachers and administrators on how to use the information gained from the survey to augment their classroom instruction plans. One teacher expressed concerns, "getting results of assessments, designing interventions and sharing with parents to eliminate or modify behaviors seems to be the most difficult task."

Despite the high level of support the assessment items received as indicated by the survey mean ratings, 7% of comments mentioned that there are too many assessments already (Fig. 15). The comments showed that teachers did not understate the value of assessments but rather mentioned their concern that teachers are not being given enough time to teach, and that many of the assessments seem to be testing for similar areas of development and are not all necessary.

Finally 3% of the comments mentioned the value of knowing whether or not a child attended pre-school and how they performed in that environment (Fig. 15). Some comments mentioned the need for making pre-school available to more students and others discussed the possibility of opening lines of communication between kindergarten teachers and pre-school teachers. One teacher noted the necessity of pre-school, "The bar has been raised so high for these five and six-year olds that it is becoming harder and harder for those children who have never been exposed to a structured academic environment."

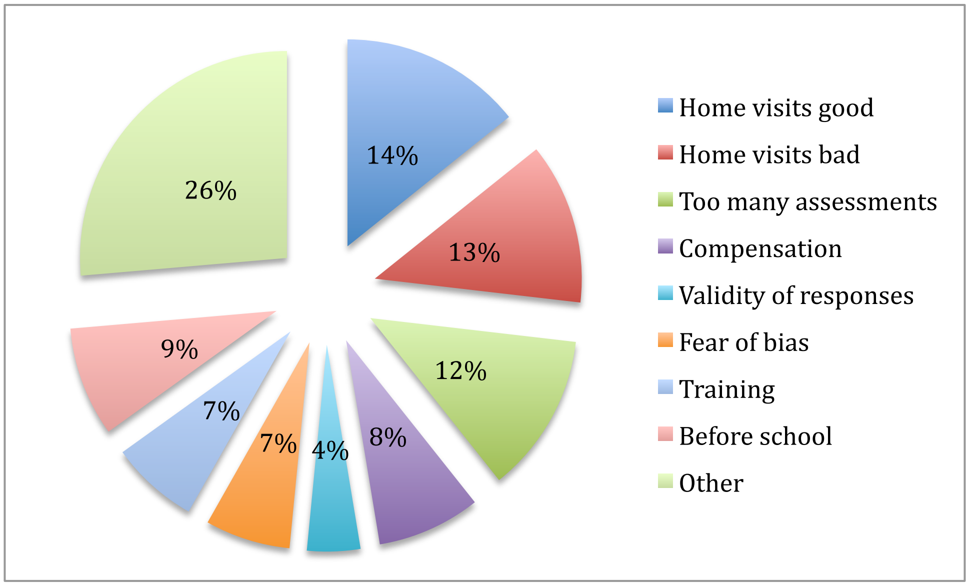


Figure 16: Breakdown of comments regarding instrument

At the end of the second part of the survey, the respondents were again offered a space to include their opinions or suggestions regarding the instrument and delivery of the assessment. In this section, 842 respondents offered comments (Table 14.2). Of these, 26% were placed in the "other" category because they did not offer any new information that expanded off of the basic ratings.

The largest groups of new information were those who expressed the belief that home visits were extremely beneficial to teachers and should be required. These opinions comprised 14% of the comments in this section and discussed the necessity of understanding a child's background to inform their classroom instruction (Fig. 16). Some mentioned the relevance of "home culture" in a child's education and other's felt it can help keep parents engaged in their child's academic life. One teacher mentioned, "Administering an assessment in familiar surroundings could help with anxiety issues." Several of these comments also included the possibility of a translator being necessary for home visits.

However, 13% of comments mentioned the dangers of conducting home visits and did not condone them (Fig. 16). Most expressed a fear of going into homes in lower-income neighborhoods known for violence and other crimes. Some also mentioned that home visits would be intrusive and parents would not welcome the breech in privacy of their homes. Still others noted that some parents of lower socio-economic backgrounds would be ashamed of their dwellings and would not feel comfortable with an educator visiting them. One strong opinion mentioned, "It is not safe for teachers to visit the homes of students. I would resign if that were required of me." Another mentioned, "If we conducted home visits, we would not see how children behave away from their parents or how they interact with other children."

Again, in this section of comments, 12% mentioned that there already existed too many assessments and that adding more would be a hindrance to a child's education and a teacher's ability to address the concerns elucidated by assessment results (Fig. 16). Similarly to the first comment section, these concerns were raised about redundancy of the areas these assessments evaluate. One teacher mentioned the negative impact of too many assessments on children, "constant assessing is stressful on young children." Another said, "I don't think adding to the assessments that are already in place for kindergarten will be helpful. It will just take away instructional time. What is really needed in the primary grade is more academic support staff."

Similar to the previous comment section, 9% of comments mentioned the need to administer this assessment before children entered school, for many of the same reasons as previously mentioned (Fig. 16). One teacher suggested that "Assessments should be used for making balanced classes and student placement. Also, they should be used as a jumping off point, understanding that kindergarten students don't always perform their best on initial assessments due to nerves." This comment brings to light the possibility of multiple assessments being able to garner a more accurate picture of a child's development.

In this section, a solid 8% of the comments noted the need for compensation before asking teachers to conduct these assessments (Fig. 16). Teachers felt that they were already underpaid and overburdened and that they could not be expected to carry out an additional responsibility during or before the school year without adequate compensation. Each of these comments stated an explicit need for needing extra pay to complete this task.

In this section, 7% of the comments also mentioned the need for training for teachers (Fig. 16). Some mentioned the need for teachers to be trained before entering the homes of children and others showed the importance of teaching educators how to use the assessment information. One teacher brought up the issue of standardizing the impact of this assessment by ensuring that all teachers properly utilize it, "I think that teachers would need to be educated on how to use the information not for biases, but for meeting the needs of the child."

The next category of comments addressed explicitly the fear of forming a bias against the child based on the assessment information or by conducting home visits. This group comprised 7% of the total comments in this section (Fig. 16). Some teachers noted that forming a bias against a child based off of their background is not ideal but sometimes hard to control. One respondent summarized this sentiment, "As an educator, it is hard not to form biases based on initial assessing of students, but it's also easy to look past those and do what's best for the child."

Finally, 4% of the comments in this section regarded doubts some teachers had about the validity of the responses (Fig 16). Some teachers mentioned the limitations of an assessment conducted once, due to "off" days a child might have or anxiety caused by shyness or being in a new environment. Others mentioned that home visits could also skew the results if a child feels pressured from the presence of parents or siblings to respond in a particular way, that may not be accurate or reflective of their feelings. One teacher mentioned that "parents would try to answer for children" in the case of home assessments.

Discussion

Major Findings

Language Development was the highest ranked domain for kindergarten assessment by teachers. Immediately following Language Development came the domain of Cognitive Development. Both domains were rated highly on an absolute scale as well relative to the other domains and are thus deemed extremely important by kindergarten teachers to assess children on prior to entering kindergarten. Briefly looking at the comments section shows that one reason for this continued importance on literacy and cognition could be the new higher standards put in place by the Common Core and the rigorous pace of kindergarten that many students seem to find challenging.

The lowest domain of assessment was Approaches to Learning. In terms of the actual assessment, teachers reacted on the notion of the comprehensive inclusion of the five domains favorably. Almost all the items received a rate over 3.5, considered neutral on a 1 to 6 scale, with the exception of parent's education level. Additionally, each of the ratings for the domains was over 3.5, the neutral score. This suggests that teachers do see the value in addressing Emotional/Social Development, Language Development, Cognitive Development, Health and Physical Development and Approaches to Learning. Further, teachers did not seem to believe that it was likely that the assessment information would be harmful or prevent them from meeting a child's needs.

On the whole, teachers also expressed their belief that the information gained by the assessment would be helpful in their day-to-day work in the classroom, with a score of 5.06 on a 1 to 6 scale, 1 being least helpful and 6 being most helpful. However, fewer saw the value in conducting multiple assessments a year. The suggestion of multiple assessments only received a rating of 4.29 on a 1 to 6 scale, 1 being least helpful and 6 being most helpful. The comments may provide some insight into the reasons why: many teachers mentioned that they felt there were too many assessments already in existence and that adding more would cause redundancy in the areas being evaluated. Many comments raised concerns regarding the time these assessments take away from classroom instruction.

There was overwhelming support for the notion of using a tablet to conduct these assessments. However, there was significant difference in the rating of comfort using a tablet among teachers by groups of years teaching. As teachers had more experience in the classroom, presumably the older teachers, they seemed less comfortable with the prospect of administering these assessments on a tablet. It is possible that teachers might need training to be able to confidently use new technology to deliver the assessment.

However, on the whole, teachers did not endorse the idea of conducting the assessment during home visits. This proposal received a rate of 3.33, less than the neutral score of 3.5, on a 1 to 6 scale, 1 being least support and 6 being most support. The comments indicated that while many teachers saw the value in understanding a child's background and culture, many teachers also felt that home visits would be dangerous and intrusive. They did not feel safe entering the homes of families living in low income neighborhoods and also felt it could seem disrespectful to the families to infringe in their lives without asking permission first.

Further, while this was not examined quantitatively within the survey, the comments indicated that several teachers stated they would require compensation to carry out an additional task, such administering this assessment to all their students, especially if home visits would be required.

Limitations

Because of the low response rate, it is impossible to state conclusively that the findings are representative of the population, being all of the kindergarten teachers in North Carolina public schools. Our results found discrepancies in endorsement of different items within a domain. Teachers rated some items within a given domain much higher or lower than the other items, which could mean that these domains include diverse topics. Given the informal nature of the comments, the coding system used to evaluate them and quantify the rate at which certain ideas were mentioned was not perfectly accurate. Further, some comments fell into more than one of the categories and so the percentages are not entirely representative of frequency. However, given the large sample size, it is our belief that the results of the study remain valid and informative.

Major recommendations

The think tank should take several key pieces of information from this study to inform future decisions. First, it is imperative that the assessment include, in some way, nearly all of the 22 items on the survey, with the possible exception of parent's education level. Additional items to consider for inclusion in the assessment would be information about a child's potential pre-school experience as well as a parent's level of support or willingness to engage in a child's education.

Further, the manner of assessment must be created so that the evaluation of social and emotional development is objective and not liable to influence by a child's anxiety, parental interruption, or a teacher's biases.

Further, teachers should be trained on how to administer this assessment, especially if a tablet is to be used or if it is conducted in the home.

Home visits should be considered carefully, given the reservations teachers have about their safety. Teachers should be provided with a translator or simply a second individual to accompany them for their personal safety.

Further, both teachers and educators should be trained how to use the information gained from the assessment to address unique needs in children and to generally inform instruction. This training should be standardized so that all public schools are utilizing the information fairly and effectively. Further, the training can help ameliorate the fear some of the information gained from an assessment can cause biases that may be harmful in the instruction of a child.

Acknowledgements

I would like to offer my sincerest thanks to Professor Ken Dodge for his patience, guidance, and faith in me throughout this process. Without his support and assistance, this project would not have been possible. I would also like to extend my gratitude to the North Carolina Think Tank for their general willingness to involve me in their work, to Ryan Denniston with Duke SSRI for helping me navigate SPSS, as well as the Public Policy department, without which I would never have discovered the power of bridging the research and policy divide.

Finally I would like to thank my friends and family for bearing with me throughout this stressful endeavor and serving as steadfast motivation and encouragement to do my best.

References

Boyer, E.L. (1991). *Ready to Learn: A Mandate for the Nation*. Princeton, N.J.: Carnegie Foundation for the Advancement of Teaching. Lawrenceville, N.J.: Princeton University Press.

National Research Council. *Early Childhood Assessment: Why, What, and How*. Washington, DC: The National Academies Press, 2008.

Gillard, A. M. (2008). *The predictive validity of kindergarten assessment.* Loyola University Chicago). *ProQuest Dissertations and Theses,* , 124-n/a. Retrieved from http://search.proquest.com/docview/304556506?accountid=10598

Heaviside, S., & Farris, S. (1993). *Public school kindergarten teachers' views on children's readiness for school.* Washington, DC: National Center for Education Statistics.

Kagan, S.L., E. Moore, & S. Bredekamp, eds. 1995. Reconsidering children's early development and learning: Toward common views and vocabulary. Washington, DC: National Education Goals Panel. Online: www.negp.gov/Reports/child-ea.htm

La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. Review of Educational Research, 70, 443-484. Note: the correlation between school readiness screenings and children's later academic and cognitive competency was .49, considered to be a moderate effect size. The correlation between these screenings and children's subsequent social competency was .27, considered to be a small effect size. The percentages cited are the result of squaring these effect sizes.

Lee, V., & Burkham, D. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. Washington, DC: Economic Policy Institute.

Lin, H.-L., Lawrence, F. R., & Gorrell, J. (2003). *Kindergarten teachers' views of children's readiness for school*. Early Childhood Research Quarterly, Volume 18, Issue 2, 225-237. Meisels.

"NC RACE TO THE TOP." *NC Race to the Top*. Department of Public Instruction, 10 Jan. 2010. Web. 10 Mar. 2013.

N.C. Gen. Stat. § 115C-83.1E. Developmental screening and kindergarten entry assessment.

N.C. Gen. Stat. § 115C-174.11. Components of the testing program

"North Carolina General Assembly - General Statutes." *North Carolina General Assembly - General Statutes*. N.p., n.d. Web. 10 Mar. 2013.

National Association of State Boards of Education [NASBE]. (1991). *Caring communities: Supporting young children and families.* The Report of the National Task Force on School Readiness. Alexandria, VA: Author.

Pianta, R. C., & McCoy, S. J.(1997). The first day ofschool: *The predictive validity of early school screening*. Journal of Applied Developmental Psychology, 18, 1-22.

C. Scott-Little and K. Maxwell, School Readiness in North Carolina: Strategies for Defi ning, Measuring, and Promoting Success for All Children (Greensboro, N.C.: UNCG Regional Educational Laboratory at SERVE, 2000).

Stedron, J. (2009). Technical report: State approaches to school readiness assessment. Denver, CO:National Conference of State Legislators.

Wright, C., Diener, M., & Kay, S. C. (2000). *School readiness of low-income children at risk for school failure.* Journal of Children & Poverty, 6, 99-117.

Zill, N., & J. West. 2001. *Entering kindergarten: A portrait of American children when they begin school.* Findings from the condition of education 2000. NCES #2001-035. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
Online: http://nces.ed.gov/pubs2001/2001035.pdf.

Tables

Table 1: Breakdown of sample by school	Total Count
Alamance Burlington	37
Alexander County	5
Alleghany County	3
Alpha Academy	1
American Renaissance School	1
Anson County	9
Ashe County	3
Asheboro City Schools	9
Asheville City Schools	10
Avery County Schools	8
Beaufort County Schools	18
Bertie County Schools	5
Bladen County Schools	6
Brunswick County Schools	11
Buncombe County Schools	58
Burke County Schools	21
Cabarrus County Schools	33
Caldwell County Schools	23
Camden County Schools	3
Carter Community Charter	1
Carteret County Schools	11
Caswell County Schools	3
Catawba County Schools	19
Chapel Hill-Carrboro City Schools	17
Charlotte Mecklenberg Schools	266
Charlotte Mecklenburg Schools	20
Charter Day Schools	18
Chatham County Schools	12
Cherokee County Schools	7
Clay County Schools	2
Cleveland County Schools	17
Clinton City Schools	2
Clover Garden Schools	1
Columbus Charter Schools	1
Columbus County Schools	12
Craven County Schools	13
Cumberland County Schools	87
Currituck County Schools	8
Dare County Schools	13
Davidson County Schools	38
Davie County Schools	14
Duplin County Schools	14
Durham Public Schools	33
Edgecombe County Schools	11

Elizabeth City-Pasquotank Public Schools	5
Elkin City Schools	1
Forsyth County Schools	86
Franklin County Schools	14
Gaston County Schools	51
Gates County Schools	5
Granville County Schools	14
Greene County Schools	5
Greensboro Academy	1
Guilford County Schools	146
Halifax County Schools	3
Harnett County Schools	23
Haywood County Schools	24
Henderson Collegiate	2
Henderson County Public Schools	26
Hertford County Schools	5
Hickory City Schools	5
Hoke County Schools	26
Hyde County Schools	1
Iredell-Statesville Schools	23
Jackson County Schools	12
Johnston County Schools	69
Jones County Schools	4
Kannapolis City Schools	10
Lee County Schools	26
Lenoir County Schools	24
Lincoln County Schools	25
Macon County Schools	9
Madison County	2
Martin County Schools	5
McDowell County Schools	17
Mitchell County	7
Montgomery County Schools	2
Moore County Schools	9
Mooresville Graded Schools	6
Mountain Island Charter School	3
Nash-Rocky Mount Schools	40
New Hanover County Schools	70
Newton Conover City Schools	3
Northampton County Schools	3
Onslow County Schools	7
Orange County Schools	28
Pamlico County Schools	2
Pender County Schools	19
Perquimans County Schools	4
Person County Schools	6
Piedmont Community Charter	3
Pitt County Schools	42
Polk County Schools	6

Public Schools of Robeson County	27
Quality Education Academy	2
Queen's Grant Community	1
Randolph County Schools	37
Richmond County Schools	9
Roanoke Rapids Graded Schools	7
Rockingham County Schools	22
Rowan Salisbury Schools	31
Rutherford County Schools	21
Sampson County Schools	14
Scotland County Schools	7
Socrates Academy	2
Stanly County Schools	20
Stokes County Schools	13
Surry County Schools	22
Swain County	3
The Learning Center	1
Transylvania County Schools	9
Two Rivers Community School	1
Tyrrell County Schools	3
Union County Public Schools	92
Vance County Schools	6
Wake County Schools	271
Warren County Schools	6
Washington County Schools	5
Watauga County Schools	15
Wayne County Public Schools	33
Wilkes County Schools	27
Wilson County Schools	17
Yadkin County Schools	12
Yancey County Schools	4
Grand Total	**2573**

Table 2.1: Mean rating of importance for suggested assessment items

	N	Minimum	Maximum	Mean	Std. Deviation
Receptive Language	2432	1	6	5.36	.835
Expressive Language	2432	1	6	5.23	.911
Developmental Concerns	2432	1	6	5.23	1.118
Literacy	2432	1	6	5.18	1.060
Compliance	2432	1	6	5.12	1.026
Math Skills	2432	1	6	5.00	1.117
Social Norms	2432	1	6	4.96	1.067
Self-help	2432	1	6	4.96	1.241
Fine Motor Skills	2432	1	6	4.84	1.067
Aggression	2432	1	6	4.82	1.359
Cooperate with children	2432	1	6	4.81	1.107
Self Regulate	2432	1	6	4.59	1.251
Health of Child	2432	1	6	4.53	1.267
Gross Motor Skills	2432	1	6	4.50	1.236
Conflict Resolution	2432	1	6	4.47	1.220
Family Stability	2432	1	6	4.32	1.332
Motivation to Attend School	2432	1	6	4.30	1.462
Intellectual Curiosity	2432	1	6	4.17	1.308
Anxiety/Sadness	2432	1	6	4.16	1.314
Developmental History	2432	1	6	4.08	1.550
Family Health History	2432	1	6	3.57	1.452
Parent’s Education Level	2432	1	6	3.08	1.389

Table 2.2: Mean rating of importance for domains of assessment

	N	Minimum	Maximum	Mean	Std. Dev.
Language Development and Communication	2432	1.00	6.00	5.2562	.73410
Cognitive Development	2432	1.00	6.00	5.1147	.84590
Emotional-Social Development	2432	1.00	6.00	4.7350	.91079
Health and Physical Development	2432	1.00	6.00	4.3084	.96055
Approaches to Learning	2432	1.00	6.00	3.8496	1.16542

Table 2.3: Mean response for questions regarding survey delivery

	N	Minimum	Maximum	Mean	Std. Dev.
Support for using a tablet	2409	1	6	5.17	1.238
Belief that assessment information is helpful	2407	1	6	5.06	1.111
Value for multiple assessments per year	2407	1	6	4.29	1.506
Comfort administering assessment in the home	2409	1	6	3.33	1.808
Belief that information gained could be harmful for instruction	2406	1	6	2.35	1.312

Table 3.1: Frequency of responses for support of using tablet

		Frequency	Valid Percent
Rating	1 (least support)	62	2.6
	2	67	2.8
	3	142	5.9
	4	230	9.5
	5	541	22.5
	6 (most support)	1367	56.7
	Total	2409	100.0

Table 3.2: Frequency of responses for support of administering assessment in the home

		Frequency	Valid Percent
Rating	1 (least support)	584	24.2
	2	339	14.1
	3	356	14.8
	4	358	14.9
	5	370	15.4
	6 (most support)	402	16.7
	Total	2409	100.0

Table 3.3: Frequency of responses for rate of helpfulness of assessments

		Frequency	Valid Percent
Rating	1 (least helpful)	22	.9
	2	55	2.3
	3	176	7.3
	4	342	14.2
	5	725	30.1
	6 (most helpful)	1087	45.2
	Total	2407	100.0

Table 3.4 Frequency of responses for rate of helpfulness of multiple assessments per year

		Frequency	Valid Percent
Rating	1 (least helpful)	163	6.8
	2	187	7.8
	3	321	13.3
	4	484	20.1
	5	631	26.2
	6 (most helpful)	621	25.8
	Total	2407	100.0

Table 3.5: Frequency of responses for likeliness that assessment information would be harmful for instruction

		Frequency	Valid Percent
Rating	1 (least likely)	771	32.0
	2	729	30.3
	3	461	19.2
	4	257	10.7
	5	120	5.0
	6 (most likely)	68	2.8
	Total	2406	100.0

Table 4.1: Significant ANOVA results for comfort conducting home visits among locale sizes

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	51.660	3	17.220	5.299	.001
Within Groups	7815.658	2405	3.250		
Total	7867.318	2408			

Table 4.2: Comparing mean comfort with home visits between Group 1 and 2

Locale Size	N	Mean comfort with home visits	Std. Deviation
1	515	3.18	1.768
2	777	3.53	1.846

Table 4.3: T-test equality of means between Group 1 and 2

	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Equal variances assumed	-3.339	1290	.001	-.344	.103	-.547	-.142
Equal variances not assumed	-3.368	1133.243	.001	-.344	.102	-.545	-.144

Table 4.4: Comparing mean comfort with home visits between Group 2 and 3

Locale size	N	Mean comfort with home visits	Std. Deviation
2	777	3.53	1.846
3	1074	3.25	1.788

	Table 4.5: T-test equality of means between Group 2 and 3						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Equal variances assumed	3.294	1849	.001	.281	.085	.114	.449
Equal variances not assumed	3.277	1640.388	.001	.281	.086	.113	.450

Table 5.1: Significant ANOVA results for assessment questions across years teaching

		df	Mean Square	F	Sig.
Home visits	Between Groups	3	22.689	6.996	.000
	Within Groups	2405	3.243		
	Total	2408			
Multiple assessments	Between Groups	3	12.910	5.729	.001
	Within Groups	2403	2.253		
	Total	2406			
Tablet	Between Groups	3	11.412	7.503	.000
	Within Groups	2405	1.521		
	Total	2408			

Table 5.2: Comparing mean responses between years teaching

Years		Multiple assessments	Tablets	Home visits
1	Mean	4.47	5.30	3.52
	N	656	656	656
	Std. Deviation	1.364	1.134	1.729
2	Mean	4.29	5.25	3.40
	N	523	524	524
	Std. Deviation	1.503	1.133	1.843
3	Mean	4.23	5.11	3.10
	N	757	758	758
	Std. Deviation	1.551	1.311	1.820
4	Mean	4.12	4.99	3.35
	N	471	471	471
	Std. Deviation	1.596	1.339	1.821
Total	Mean	4.29	5.17	3.33
	N	2407	2409	2409
	Std. Deviation	1.506	1.238	1.808

Table 5.3: Correlation between years teaching and comfort using tablet

		Years	Tablet
Years	Pearson Correlation	1	-.095**
	Sig. (2-tailed)		.000
	N	2584	2409
Tablet	Pearson Correlation	-.095**	1
	Sig. (2-tailed)	.000	
	N	2409	2409

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5.4: Correlation between years teaching and comfort conducting home visits

		Years	Home visits
Years	Pearson Correlation	1	-.059**
	Sig. (2-tailed)		.004
	N	2584	2409
Home visits	Pearson Correlation	-.059**	1
	Sig. (2-tailed)	.004	
	N	2409	2409

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5.5: Correlation between years teaching and rating of helpfulness of multiple assessments

		Years	Multiple assessments
Years	Pearson Correlation	1	-.083**
	Sig. (2-tailed)		.000
	N	2584	2407
Multiple assessments	Pearson Correlation	-.083**	1
	Sig. (2-tailed)	.000	
	N	2407	2407

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6.1: Significant ANOVA results for assessment questions across regions

		df	Mean Square	F	Sig.
Home visits	Between Groups	8	16.750	5.198	.000
	Within Groups	2400	3.222		
	Total	2408			
Multiple assessments	Between Groups	8	6.653	2.954	.003
	Within Groups	2398	2.252		
	Total	2406			

Table 6.2: Comparing mean responses among regions

Region		Multiple assessments	Home visits
North Central	Mean	4.32	3.12
	N	460	460
	Std. Deviation	1.452	1.765
Northeast	Mean	4.37	3.76
	N	115	115
	Std. Deviation	1.512	1.843
Northwest	Mean	4.15	3.04
	N	230	230
	Std. Deviation	1.616	1.688
Piedmont-Triad/Central	Mean	4.16	3.44
	N	418	418
	Std. Deviation	1.547	1.855
Sandhills/South Central	Mean	4.60	3.84
	N	214	214
	Std. Deviation	1.403	1.788
Southeast	Mean	4.36	3.41
	N	216	217
	Std. Deviation	1.484	1.798
Southwest	Mean	4.31	3.21
	N	504	504
	Std. Deviation	1.457	1.799
Unspecified	Mean	4.66	3.60
	N	44	45
	Std. Deviation	1.346	1.724
Western	Mean	4.04	3.30
	N	206	206
	Std. Deviation	1.621	1.836
Total	Mean	4.29	3.33
	N	2407	2409
	Std. Deviation	1.506	1.808

Table 6.3: Correlation between rating of helpfulness of multiple assessments and comfort conducting home visits

		Multiple assessments	Home visits
Multiple assessments	Pearson Correlation	1	.385**
	Sig. (2-tailed)		.000
	N	2407	2407
Home visits	Pearson Correlation	.385**	1
	Sig. (2-tailed)	.000	
	N	2407	2409

**. Correlation is significant at the 0.01 level (2-tailed).

Table 7.1: Significant ANOVA results for assessment items across regions

		Sum of Squares	df	Mean Square	F	Sig.
Gross motor	Between Groups	35.333	8	4.417	2.909	.003
	Within Groups	3678.618	2423	1.518		
	Total	3713.950	2431			
Child health	Between Groups	38.615	8	4.827	3.025	.002
	Within Groups	3866.552	2423	1.596		
	Total	3905.167	2431			
Intellectual curiosity	Between Groups	37.777	8	4.722	2.777	.005
	Within Groups	4120.445	2423	1.701		
	Total	4158.222	2431			
Parent education	Between Groups	65.273	8	8.159	4.275	.000
	Within Groups	4623.949	2423	1.908		
	Total	4689.222	2431			

Table 7.2: Significant variation of means for assessment items across regions

Region		Gross motor	Intellectual curiosity	Child health	Parent education
North Central	Mean	4.31	4.10	4.46	3.03
	N	466	466	466	466
	Std. Deviation	1.323	1.318	1.316	1.422
Northeast	Mean	4.67	4.19	4.58	3.40
	N	115	115	115	115
	Std. Deviation	1.114	1.420	1.228	1.362
Northwest	Mean	4.55	4.15	4.68	3.02
	N	231	231	231	231
	Std. Deviation	1.197	1.367	1.194	1.332
Piedmont-Triad/Central	Mean	4.51	4.26	4.59	3.20
	N	421	421	421	421
	Std. Deviation	1.251	1.269	1.228	1.374
Sandhills/South Central	Mean	4.70	4.44	4.75	3.44
	N	217	217	217	217
	Std. Deviation	1.225	1.301	1.195	1.390
Southeast	Mean	4.58	4.21	4.58	3.04
	N	221	221	221	221
	Std. Deviation	1.243	1.240	1.317	1.487
Southwest	Mean	4.48	4.03	4.35	2.90
	N	509	509	509	509
	Std. Deviation	1.226	1.304	1.291	1.341
Unspecified	Mean	4.83	4.48	4.74	3.02
	N	46	46	46	46
	Std. Deviation	.926	1.410	1.307	1.483
Western	Mean	4.51	4.07	4.51	3.01
	N	206	206	206	206
	Std. Deviation	1.151	1.249	1.237	1.318
Total	Mean	4.50	4.17	4.53	3.08
	N	2432	2432	2432	2432
	Std. Deviation	1.236	1.308	1.267	1.389

Table 8.1: Significant ANOVA results for assessment items across locale sizes

		Sum of Squares	df	Mean Square	F	Sig.
Self regulate	Between Groups	20.894	3	6.965	4.467	.004
	Within Groups	3785.316	2428	1.559		
	Total	3806.211	2431			
Resolve conflicts	Between Groups	19.232	3	6.411	4.326	.005
	Within Groups	3598.003	2428	1.482		
	Total	3617.235	2431			
Fine motor	Between Groups	32.855	3	10.952	9.715	.000
	Within Groups	2737.144	2428	1.127		
	Total	2769.998	2431			
Gross motor	Between Groups	84.222	3	28.074	18.779	.000
	Within Groups	3629.728	2428	1.495		
	Total	3713.950	2431			
Expressive language	Between Groups	17.993	3	5.998	7.280	.000
	Within Groups	2000.282	2428	.824		
	Total	2018.275	2431			
Receptive language	Between Groups	14.449	3	4.816	6.954	.000
	Within Groups	1681.738	2428	.693		
	Total	1696.187	2431			
Aggression	Between Groups	35.693	3	11.898	6.481	.000
	Within Groups	4457.247	2428	1.836		
	Total	4492.941	2431			
Self help	Between Groups	37.458	3	12.486	8.185	.000
	Within Groups	3703.673	2428	1.525		
	Total	3741.131	2431			

Motivation school	Between Groups	26.985	3	8.995	4.225	.005
	Within Groups	5168.692	2428	2.129		
	Total	5195.677	2431			
Intellectual curiosity	Between Groups	36.136	3	12.045	7.095	.000
	Within Groups	4122.086	2428	1.698		
	Total	4158.222	2431			
Child health	Between Groups	57.245	3	19.082	12.040	.000
	Within Groups	3847.922	2428	1.585		
	Total	3905.167	2431			
Parent education	Between Groups	53.784	3	17.928	9.391	.000
	Within Groups	4635.438	2428	1.909		
	Total	4689.222	2431			
Family health history	Between Groups	28.246	3	9.415	4.482	.004
	Within Groups	5100.424	2428	2.101		
	Total	5128.669	2431			

Table 8.2: Significant variation of means for assessment items across locale sizes

Locale size		Self regulate	Resolve conflicts	Fine motor	Gross motor	Expressive language	Receptive language
1	Mean	4.47	4.32	4.66	4.19	5.12	5.24
	N	521	521	521	521	521	521
	Std. Deviation	1.316	1.275	1.118	1.314	1.009	.890
2	Mean	4.68	4.53	4.80	4.48	5.18	5.34
	N	784	784	784	784	784	784
	Std. Deviation	1.230	1.209	1.061	1.223	.887	.820
3	Mean	4.55	4.47	4.96	4.66	5.32	5.43
	N	1083	1083	1083	1083	1083	1083
	Std. Deviation	1.240	1.205	1.043	1.186	.881	.818
4	Mean	4.95	4.80	4.91	4.82	5.36	5.32
	N	44	44	44	44	44	44
	Std. Deviation	.939	.930	.802	.947	.613	.674
Total	Mean	4.59	4.47	4.84	4.50	5.23	5.36
	N	2432	2432	2432	2432	2432	2432
	Std. Deviation	1.251	1.220	1.067	1.236	.911	.835

Table 8.2 (cont.) : Significant variation of means for assessment items across locale sizes

Locale size		Aggression	Self help	Motivation school	Intellectual curiosity	Child health	Parent education	Family health history
1	Mean	4.65	4.74	4.10	3.96	4.25	2.82	3.39
	N	521	521	521	521	521	521	521
	Std. Deviation	1.444	1.358	1.474	1.296	1.370	1.352	1.444
2	Mean	4.78	4.96	4.32	4.16	4.55	3.11	3.57
	N	784	784	784	784	784	784	784
	Std. Deviation	1.415	1.242	1.469	1.307	1.276	1.379	1.458
3	Mean	4.90	5.06	4.37	4.26	4.65	3.20	3.66
	N	1083	1083	1083	1083	1083	1083	1083
	Std. Deviation	1.283	1.178	1.447	1.298	1.185	1.396	1.444
4	Mean	5.36	5.14	4.36	4.45	4.75	2.93	3.34
	N	44	44	44	44	44	44	44
	Std. Deviation	.750	.905	1.399	1.422	1.332	1.437	1.493
Total	Mean	4.82	4.96	4.30	4.17	4.53	3.08	3.57
	N	2432	2432	2432	2432	2432	2432	2432
	Std. Deviation	1.359	1.241	1.462	1.308	1.267	1.389	1.452

Table 9.1: Significant ANOVA results for assessment item across degrees

Approaches

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	24.624	3	8.208	3.854	.009
Within Groups	5171.054	2428	2.130		
Total	5195.677	2431			

Table 9.2: Variation across degrees for rate of importance of assessment item across degrees

Approaches

Degree	Mean	N	Std. Deviation
1	4.37	1637	1.429
2	4.15	786	1.519
3	4.50	4	.577
4	4.20	5	1.924
Total	4.30	2432	1.462

Table 10.1: Significant ANOVA results for assessment items across years teaching

		df	Mean Square	F	Sig.
Self regulate	Between Groups	3	11.184	7.198	.000
	Within Groups	2428	1.554		
	Total	2431			
Resolve conflicts	Between Groups	3	11.294	7.652	.000
	Within Groups	2428	1.476		
	Total	2431			
Gross motor	Between Groups	3	6.985	4.592	.003
	Within Groups	2428	1.521		
	Total	2431			
Expressive language	Between Groups	3	3.830	4.634	.003
	Within Groups	2428	.827		
	Total	2431			
Receptive language	Between Groups	3	6.892	9.988	.000
	Within Groups	2428	.690		
	Total	2431			
Literacy	Between Groups	3	5.572	4.985	.002
	Within Groups	2428	1.118		
	Total	2431			
Aggression	Between Groups	3	11.914	6.490	.000
	Within Groups	2428	1.836		
	Total	2431			
Compliance	Between Groups	3	4.814	4.598	.003
	Within Groups	2428	1.047		
	Total	2431			
Social norms	Between Groups	3	5.389	4.755	.003
	Within Groups	2428	1.134		
	Total	2431			
Self help	Between Groups	3	8.892	5.812	.001
	Within Groups	2428	1.530		
	Total	2431			
Motivation school	Between Groups	3	11.341	5.335	.001
	Within Groups	2428	2.126		
	Total	2431			
Intellectual curiosity	Between Groups	3	10.488	6.171	.000
	Within Groups	2428	1.700		
	Total	2431			
Anxiety	Between Groups	3	7.777	4.521	.004
	Within Groups	2428	1.720		
	Total	2431			
	Between Groups	3	7.153	5.754	.001

Developmental concerns	Between Groups	3	7.153	5.754	.001
	Within Groups	2428	1.243		
	Total	2431			
Child health	Between Groups	3	13.441	8.444	.000
	Within Groups	2428	1.592		
	Total	2431			
Parent education	Between Groups	3	22.374	11.753	.000
	Within Groups	2428	1.904		
	Total	2431			
Family health history	Between Groups	3	10.197	4.857	.002
	Within Groups	2428	2.100		
	Total	2431			

Table 10.2: Significant variation of means for assessment items across years teaching

Years teaching		Self regulate	Resolve conflicts	Gross motor	Expressive language	Receptive language	Literacy	Aggression	Compliance
1	Mean	4.56	4.45	4.65	5.15	5.27	5.25	4.73	5.07
	N	660	660	660	660	660	660	660	660
	Std. Deviation	1.209	1.163	1.176	.949	.864	1.048	1.277	.996
2	Mean	4.45	4.32	4.45	5.18	5.27	5.21	4.68	5.02
	N	531	531	531	531	531	531	531	531
	Std. Deviation	1.272	1.245	1.233	.954	.881	1.043	1.409	1.081
3	Mean	4.56	4.44	4.42	5.28	5.42	5.19	4.87	5.17
	N	764	764	764	764	764	764	764	764
	Std. Deviation	1.313	1.285	1.298	.888	.814	1.056	1.398	1.017
4	Mean	4.81	4.68	4.49	5.33	5.49	5.02	5.01	5.23
	N	477	477	477	477	477	477	477	477
	Std. Deviation	1.156	1.132	1.205	.831	.749	1.088	1.327	1.004
Total	Mean	4.59	4.47	4.50	5.23	5.36	5.18	4.82	5.12
	N	2432	2432	2432	2432	2432	2432	2432	2432
	Std. Deviation	1.251	1.220	1.236	.911	.835	1.060	1.359	1.026

Table 10.3 (cont.): Significant variation of means for assessment items across years teaching

Years		Social norms	Self help	Motivation school	Intellectual curiosity	Anxiety	Developmental concerns	Child health	Parent education	Family health history
1	Mean	4.93	4.85	4.26	4.10	4.16	5.11	4.55	3.08	3.48
	N	660	660	660	660	660	660	660	660	660
	Std. Deviation	1.051	1.272	1.431	1.323	1.280	1.180	1.261	1.433	1.464
2	Mean	4.84	4.90	4.22	4.06	4.05	5.18	4.37	2.90	3.46
	N	531	531	531	531	531	531	531	531	531
	Std. Deviation	1.106	1.271	1.438	1.302	1.286	1.105	1.304	1.375	1.406
3	Mean	5.00	4.98	4.23	4.16	4.13	5.27	4.50	3.02	3.60
	N	764	764	764	764	764	764	764	764	764
	Std. Deviation	1.054	1.233	1.519	1.327	1.351	1.128	1.270	1.354	1.455
4	Mean	5.08	5.14	4.54	4.38	4.34	5.37	4.76	3.39	3.76
	N	477	477	477	477	477	477	477	477	477
	Std. Deviation	1.055	1.153	1.416	1.239	1.320	1.007	1.200	1.350	1.466
Total	Mean	4.96	4.96	4.30	4.17	4.16	5.23	4.53	3.08	3.57
	N	2432	2432	2432	2432	2432	2432	2432	2432	2432
	Std. Deviation	1.067	1.241	1.462	1.308	1.314	1.118	1.267	1.389	1.452

Table 11.1: Significant ANOVA results for domains across years teaching

		df	Mean Square	F	Sig.
Approaches	Between Groups	3	13.781	10.263	.000
	Within Groups	2428	1.343		
	Total	2431			
Health	Between Groups	3	3.863	4.203	.006
	Within Groups	2428	.919		
	Total	2431			
Emotional	Between Groups	3	7.150	8.701	.000
	Within Groups	2428	.822		
	Total	2431			

Table 11.2: Significant variation of means for domains across years teaching

Years teaching		Approaches	Health	Emotional
1	Mean	3.8157	4.3354	4.6890
	N	660	660	660
	Std. Deviation	1.15391	.94833	.85751
2	Mean	3.7269	4.2159	4.6234
	N	531	531	531
	Std. Deviation	1.13646	.94032	.92898
3	Mean	3.8050	4.2797	4.7482
	N	764	764	764
	Std. Deviation	1.20406	.99479	.93650
4	Mean	4.1048	4.4200	4.9020
	N	477	477	477
	Std. Deviation	1.11551	.93376	.89819
Total	Mean	3.8496	4.3084	4.7350
	N	2432	2432	2432
	Std. Deviation	1.16542	.96055	.91079

Table 12.1: Significant ANOVA results for domains across locale sizes

		df	Mean Square	F	Sig.
Approaches	Between Groups	3	12.203	9.074	.000
	Within Groups	2428	1.345		
	Total	2431			
Health	Between Groups	3	11.221	12.332	.000
	Within Groups	2428	.910		
	Total	2431			
Emotional	Between Groups	3	5.314	6.448	.000
	Within Groups	2428	.824		
	Total	2431			
Language	Between Groups	3	2.459	4.583	.003
	Within Groups	2428	.537		
	Total	2431			
Cognitive	Between Groups	3	3.027	4.247	.005
	Within Groups	2428	.713		
	Total	2431			

Table 12.2: Significant variation of means for assessment items across locale sizes

Locale size		Approaches	Health	Language	Cognitive	Emotional
1	Mean	3.6244	4.1011	5.1817	5.0326	4.5974
	N	521	521	521	521	521
	Std. Deviation	1.14538	.99024	.75577	.84801	.96895
2	Mean	3.8639	4.3027	5.2241	5.0810	4.7500
	N	784	784	784	784	784
	Std. Deviation	1.15785	.96441	.70139	.85628	.89162
3	Mean	3.9449	4.4092	5.3143	5.1722	4.7784
	N	1083	1083	1083	1083	1083
	Std. Deviation	1.16609	.92907	.75019	.84229	.89496
4	Mean	3.9167	4.3826	5.2803	5.2727	5.0313
	N	44	44	44	44	44
	Std. Deviation	1.20104	.93010	.51811	.58523	.73686
Total	Mean	3.8496	4.3084	5.2562	5.1147	4.7350
	N	2432	2432	2432	2432	2432
	Std. Deviation	1.16542	.96055	.73410	.84590	.91079

Table 13.1: Significant ANOVA results for domains across regions

Approaches

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	41.871	8	5.234	3.890	.000
Within Groups	3259.926	2423	1.345		
Total	3301.797	2431			

Table 13.2: Significant variation of means for domains across regions

Approaches

Region	Mean	N	Std. Deviation
North Central	3.7854	466	1.19406
Northeast	4.0058	115	1.27197
Northwest	3.8153	231	1.13126
Piedmont-Triad/Central	3.9549	421	1.14056
Sandhills/South Central	4.1444	217	1.14918
Southeast	3.8492	221	1.19527
Southwest	3.7040	509	1.12268
Unspecified	3.9710	46	1.20662
Western	3.7540	206	1.14002
Total	3.8496	2432	1.16542

Table 14.1: Comments after items

Training	54
Doubt of objectivity	96
Parent Support	68
Pre-K Info	28
Asses before school	112
Too many assessments	61
High new standards	127
Other	308
Total	864

Table 14.2: Comments after assessment questions

Home visits good	120
Home visits bad	106
Too many assessments	105
Compensation	68
Validity of responses	35
Fear of bias	56
Training	58
Before school	72
Other	222
Total	842

Appendix 1

Which North Carolina school district do you work in?

What is your highest level of education?

How mane years have you been teaching?

Using a scale of 1 to 6, with 1 being less important and 6 being extremely important, please rate the importance of having a valid assessment of the following items for each child entering kindergarten for your day-to-day education planning. Please try to use all five response options.

1.) Ability to self-regulate (e.g. manage and control emotions, handle disappointment)
2.) Ability to resolve conflicts positively and peacefully
3.) Fine motor skills
4.) Gross motor skills
5.) Expressive language (e.g. verbalize and communicate)
6.) Receptive language (e.g. understand what is being said)
7.) Math skills (e.g. counting, recognizing basic shapes,)
8.) Literacy skills (e.g. concepts of print, alphabet knowledge)
9.) Tendency toward aggressive behavior (towards others and self)
10.) Ability to comply with adult request (e.g. focus, compliance, cooperation)
11.) Ability to follow social rules and routines
12.) Self-help skills (e.g. , dressing self, eating independently)
13.) Ability to interact cooperatively with other children
14.) Motivation to attend school
15.) Intellectual curiosity
16.) Anxiety/sadness (e.g. behavioral trends, habit of stress)
17.) Developmental concerns (e.g. developmental delays, autism, physical handicaps)
18.) Child's general health (basic physical and emotional well-being)
19.) Child's developmental history (e.g. age of milestones such as potty-training)
20.) Parent's education level.
21.) Family stability (e.g. traumatic family events, health concerns)
22.) Family health history (e.g. predispositions, history of certain diseases)
23.) Please use this area to offer your own comments or expand off a response to a previous question about significant predictors of kindergarten readiness.

Assess your response to the following on a 1-6 scale, 1 being least supportive and 6 being most supportive.

24.) How much would you support conducting assessments of each child in a home visit during the first several weeks of school?
25.) How much would you support the use of a tablet or other electronic device to complete assessments?

Assess your response to the following on a 1-6 scale, 1 being least helpful and 6 being most helpful.

26.) How helpful would the information gained from this assessment be to you in meeting the needs of each of your students?

27.) How helpful would it be for this student assessment to be repeated multiple times during the course of a year to evaluate student learning and guide future instruction?

Assess your response to the following on a 1-6 scale, 1 being least likely and 6 being most likely

28.) How likely is it that the information gained from this assessment could be harmful in the education of a child? (e.g. producing biases)

29.) Please use this area to offer your own comments or expand off a response to a previous question about this particular assessment, its delivery, and its implications on your teaching.

Appendix 2

Dear School Principal or Kindergarten Teacher:

The Department of Public Instruction has received funding from the Race to the Top Early Learning Challenge grant and authorization from the General Assembly to develop an assessment that will generate a Child Profile at kindergarten entry and assess student progress through third grade. This assessment, which will incorporate assessments currently being administered in the early grades, is expected to be ready for statewide implementation by the fall of 2015. A first step in the development process is to hear from kindergarten teachers about their perspectives on which domains should be assessed at kindergarten entry and how this assessment should be conducted. Kindergarten teachers are key stakeholders and their input is critical for a quality assessment.

Below is a link to a short survey that kindergarten teachers are asked to complete by March 1. The survey should take approximately 15 minutes. Responses are anonymous and confidential. Data will be compiled into a report that will be given to a group of educators responsible for developing the assessment.

PRINCIPALS ARE ASKED TO FORWARD THIS EMAIL TO ALL KINDERGARTEN TEACHERS IN THEIR SCHOOL.

https://duke.qualtrics.com/SE/?SID=SV_9Rz5JlGkoFXvzgh

If you have any questions about the survey or its use, please contact Cindy Bagwell at cindy.bagwell@dpi.nc.gov or 919-807-3710.

Thank you in advance for taking the time to complete this survey. We look forward to receiving your responses.

Appendix 3

- Approaches to Learning

30.) Motivation to attend school
31.) Intellectual curiosity
32.) Parent's education level.

- Health & Physical Development

33.) Child's general health (basic physical and emotional well-being)
34.) Child's developmental history (e.g. age of milestones such as potty-training)
35.) Fine motor skills
36.) Gross motor skills
37.) Family stability (e.g. traumatic family events, health concerns)
38.) Family health history (e.g. predispositions, history of certain diseases)

- Language Development and Communication (English Language Arts)

39.) Expressive language (e.g. verbalize and communicate)
40.) Receptive language (e.g. understand what is being said)
41.) Literacy skills (e.g. concepts of print, alphabet knowledge)

- Cognitive Development (Construction of Knowledge, Thinking and Reasoning as it relates to Science, Social Studies, Math and the Arts)

42.) Math skills (e.g. counting, recognizing basic shapes,)
43.) Developmental concerns (e.g. developmental delays, autism, physical handicaps)

- Emotional-Social Development

44.) Ability to self-regulate (e.g. manage and control emotions, handle disappointment)
45.) Ability to resolve conflicts positively and peacefully
46.) Tendency toward aggressive behavior (towards others and self)
47.) Anxiety/sadness (e.g. behavioral trends, habit of stress)
48.) Ability to comply with adult request (e.g. focus, compliance, cooperation)
49.) Ability to interact cooperatively with other children
50.) Ability to follow social rules and routines
51.) Self-help skills (e.g. , dressing self, eating independently)