A Thesis

entitled

Evaluation of a Science Language Assessment for Preschool Students

by

Andrea L. Vantaggi

Submitted to the Graduate Faculty in partial fulfillment of the requirements for the
Master of Arts Degree in Speech Language Pathology

Joan N. Kaderavek, Ph.D., Committee Chair

Lori Pakulski, Ph.D., Committee Member

Charlene M. Czerniak, Ph.D., Committee Member

Patricia Komuniecki, Ph.D., Dean
College of Graduate Studies

The University of Toledo

May 2011
Currently, there are very few assessments used to measure, aggregate, and compare young children’s science learning. The *Preschool Science Assessment Protocol (PSAP)* is an assessment instrument designed to fulfill this need. The purpose of this study was to investigate the reliability and validity of the PSAP. Specific foci included (a) investigating alternate form reliability between two matched versions of the assessment protocol, (b) investigating inter-rater reliability, and (c) investigating the validity of the PSAP by comparing children’s language performance on the PSAP and the *Clinical Evaluation of Language Fundamentals (CELF-P2)*. Children participated in two expository adult-child book reading sessions; questions posed during the book reading documented children’s ability to respond to “science language” consisting of (a) prediction, (b) WH-questions, (c) causal relationships, (d) argumentation, and (e) vocabulary.
Acknowledgements

First and foremost, I would like to thank my thesis advisor, Dr. Joan Kaderavek, for her dedication and continued support throughout this project. The time and effort you have devoted is greatly appreciated. I would also like to thank my committee members, Dr. Lori Pakulski and Dr. Charlene Czerniak. Your input and advice toward the completion of this project is greatly appreciated. Thank you!

I would like to thank All God’s Children Preschool for being so helpful and accommodating during my study. Your dedication to your students is truly admirable and they are lucky to be in such a wonderful preschool program. Thank you!

I would also like to thank my friends and classmates whose constant support and push have helped me continue writing whenever I would hit a breaking point. A special thank you to Margie Traband who has helped me immensely from day one. I would have never hit this point without you. Thank you!

Last, but not least, I would like to thank my family for their constant support regarding any crazy idea I could possibly come up with. My parents, Mark and Linda Vantaggi, will always be my strongest fans. Their belief in me has made me the determined person that I am today. My brother, Tom, has and always will be there when I need a laugh or a much needed evening out. My grandparents, Bob and Joan Drazil, also deserve a big thank you for making me believe anything is possible. Grandpa, I know you’re always watching over us and you are missed everyday. Thank you!!
# Table of Contents

Abstract.................................................................................................................. iii
Acknowledgements................................................................................................ iv
Table of Contents.................................................................................................. v
List of Tables.......................................................................................................... vi
List of Appendices................................................................................................. vii
Chapter 1: Statement of the Problem................................................................. 1
Chapter 2: Literature Review............................................................................... 3
Chapter 3: Methodology...................................................................................... 20
Chapter 4: Results............................................................................................... 28
Chapter 5: Discussion.......................................................................................... 33
References.............................................................................................................. 39
Appendices

Appendix A: Permission Forms........................................................................... 46
Appendix B: PSAP Form: *Noses at Work*......................................................... 50
Appendix C: PSAP Form: *Let’s Ride Bikes*....................................................... 53
Appendix D: Internal Revue Board Forms......................................................... 57
List of Tables

Table 1  Subjects Age in Months..................................................24
Table 2  Randomization of Testing Materials.................................27
Table 3  Individual Scores of Testing Materials.............................32
Chapter One

Statement of the Problem

Need for Science Learning in Preschool

Recent reports state that U.S. adolescents are ranked 29th in science knowledge as compared to other children from developed countries. U.S. fifteen-year-olds know less science, on average, than students from countries like Croatia, the Czech Republic, and Liechtenstein. Finland emerged at the top of 5 countries in science, according to the 2006 survey results (Programme from International Student Assessments, 2006). More recently, President Obama—in his speech on “Education to Innovate” – indicates that federal dollars will be used to bring science and math education to all U.S. children, even during the preschool years (Obama, Speeches and Remarks, November 29, 2009). The information from each of the sources underscores the importance of instituting high-quality preschool science-based curriculum as well as promoting children’s use of “science language.” Science language refers to the high-level abstract language children use for analysis and description of science activities, including their ability to use prediction and hypothesis generation.

In the past, many educators believed that children could not learn science concepts until they could read and write; consequently science is a frequently neglected part of the preschool curriculum (Nayfeld, 2009). However, research demonstrates that
exposure to science instruction in the preschool years improves children’s literacy development and facilitates language development (Andersen, Scheuer, del Puy Pérez Echeverriam, & Teubal, 2009). In more recent years, researchers have documented the relationship between preschool science instruction and language skill development. This project begins to address the issue of how to assess young children’s language ability (i.e., use of science language) in order to begin to explore how science education impacts the high-level language performance of young children.
Chapter Two

Literature Review

The review summarizes the pertinent literature and current research related to science language development and its importance during the preschool years. Additionally, the language development process in relation to science instruction and the importance of teacher interaction and question asking is discussed. During the course of the review, current barriers are reviewed as well as the language features associated with science learning in the classroom. In addition, the value of an assessment as a tool for identifying children at-risk for later language difficulties is summarized. This summary encompasses the role of assessment, intervention, and progress monitoring in science language development.

A number of assessments are examined and compared to one another, including limitations of the current assessment tools as well as the documented need for a new assessment tool. The steps involved in test construction are explored in response to this need for a novel assessment tool in the area of early science language learning. Finally, a novel assessment instrument is proposed to identify at-risk children and allow proper planning in current curriculum settings.

Science Instruction and Language Skill Development

Several research studies have been conducted regarding the relationship between preschool science instruction and language skill development. For example, French (2004) documented the effect of ScienceStart! curriculum on 162 children between the
ages of 3 to 5 years. Over a two-year period, students were divided into two groups. One group participated in a specially designed science curriculum (n = 61) while the control group continued in their normal course of study (n = 101). Children were assessed in the fall and spring of the school year to document the effects of ScienceStart! on students’ learning. The groups were matched in science knowledge at pre-test, but there was a significant difference in the two groups’ performance on the posttest, with children in the ScienceStart! classrooms receiving a mean score of 109 and those in the control classrooms receiving a mean score of 100 (French, 2004). These findings are in accord with the hypothesis that the science curriculum leads to measurable improvements in children’s science language skills.

A second study also demonstrated the facilitative effects of early science instruction. Gelman and Brenneman (2004) implemented a program called Preschool Pathways to Science (PrePS). The theoretical foundation for PrePS postulates that domain-specific information cognitively organizes into separate mental structures (Gelman & Brenneman, 2004). As a validity check for this theoretical position, researchers observed two preschool classrooms in rural New Jersey for a month. Children completed activities related to the five senses. In the experimental classroom, children were taught using the PrePS methods focusing on “hands on” learning opportunities. In contrast, children in the comparison school participated in a “business as usual” curriculum.

After some training in the PrePS program, researchers noted that children in the first group were able to observe, describe, predict, and check new knowledge. For example, researchers observed children in an activity in which they were asked to make a
prediction about what they would see inside an apple and then, after cutting the apple, confirm their hypothesis. Children’s responses were recorded and later analyzed to evaluate their response level. Results showed that children who participated in the PrePS program performed better on post-intervention assessments of science knowledge and language skills during the activity.

The studies above demonstrate that preschool children learn best in science instruction that includes hands-on participation and exploration. The concrete manipulation of science materials and the discourse that accompanies science exploration helps children form rich mental representations (French, 2004). Early science experiences enhance children’s increased math and science achievement, literacy and academic work skills (Brenneman, Stevenson-Boyd & Frede, 2009).

Science education not only exposes young children to science concepts but science activities also expand children’s higher order cognitive and language skills by requiring children to plan, describe, predict, make hypotheses, and explore principles of cause-and-effect. Science instruction creates opportunities for young children to complete an integrated process of talking about science experiments, exchanging information, asking and answering questions, reading for information, making charts and graphs, and making real-life observations and dictating results (French, 2004).

**Language Development Relating to Science Instruction**

Guidelines issued by the National Association for the Education of Young Children (NAEYC) recognize that adults play a vital role in children’s development (NAEYC, 2009). Research also has shown that young children learn best when they are involved in environments with a predictable schedule and structure. Children learn
abstract language by interacting with adults who use high-level language and who ask inferential questions. Examining data from adult-child storybook reading studies helps understand the role of inferential questions during a predictable language routine.

A study conducted by Van Kleeck, Vander Woude and Hammett (2006) documented the effect of one-on-one book-sharing intervention on thirty children ranging from 3 to 5 years of age in relation to inferential language learning. A control group receiving no intervention was compared to a treatment group receiving twice-weekly sessions in which adults read books and asked both literal and inferential questions using scripts embedded throughout the text. Groups were compared using pre- and posttest scores; data demonstrated a significant difference in child outcomes. Standard scores on the Peabody Picture Vocabulary Test – 3rd edition (a test of receptive vocabulary ability) for the treatment group demonstrated that child vocabulary improved from a pretest score mean of 80.07 to a posttest standard score mean of 90.93. In contrast, the standard mean scores of the control group remained stable ($M = 74.53$ to $M = 74.07$). These findings indicate that book sharing with literal and inferential question-asking facilitates children’s language. A hypothesis in the current study is that children’s ability to respond to literal and inferential questions can be used to document children’s use of “science language.”

A second study by Zucker, Justice, Piasta, and Kaderavek (2010) also demonstrated the extent to which preschool teachers used literal and inferential questions during classroom-based shared reading using an informational (e.g., expository text [see below]). A group of 159 four-year-old children in twenty-five preschool teachers’ classrooms were video taped using shared reading while teacher and children’s extra textual talk was analyzed. Results suggest that inferential questioning encourages
children to participate in conversation at complex, inferential levels. Importantly, informational texts (non-narrative books) provided a successful context for inferential discourse. In the current study, the author considered if informational texts could be used as part of an assessment protocol to document children’s use of “science language.”

**Importance of Question-Asking and Answering**

**Teacher practices in question asking.**

Franke et al. (2009) explored the effect of teacher questions on child outcomes. Researchers videotaped twelve children in three elementary school classrooms (children ranged in age from seven to nine) for two weeks. Researchers asked the teachers to engage students in topical discussions related to the ongoing curriculum. In two classrooms, researchers asked the teachers to use elaborative techniques when responding to students; in the control classroom, teachers were asked to conduct “business as usual.” Elaboration techniques included following up on students’ explanations and building students’ ideas.

Video and audiotapes were used to record teachers’ teaching and discussion techniques. Classroom interaction followed a standard structure: (a) the teacher posed a problem, (b) students discussed the solution within a smaller group, and (c) the teacher led a classroom discussion. Results indicated that the control classroom students answered nine questions correctly compared to children in the experimental classroom who correctly answered eighteen questions. Results supported the hypothesis that teacher support for student high-level thinking enhances student-learning outcomes.
Child question asking.

Children’s interest in exploring “everyday” science concepts begins at early ages. Preschool children exhibit a sense of inquiry by counting their steps, creating patterns, and item comparison. They also notice and question their surrounding. For example, preschool children have been noted to ask questions such as, “Where do baby cows come from?” or ”Why is that boy crying?”. Supporting and encouraging children’s question asking helps facilitate children’s language ability and higher-ordering thinking.

McKenna and Stahl (2003) report that adults should ask questions representing three levels of thinking: literal questions, inferential questions (questions that cannot be answered from items or objects in the environment), and critical questions (questions reflecting a person’s value system).

A study conducted by Yaden, Smolkin, and Conlon (1989) documented preschoolers’ question types during home storybook reading sessions. Nine parent-child dyads were repeatedly recorded and analyzed over a two-year period. In the first study, 2 male children were recorded weekly during book reading sessions. Over the two-year period, the boys asked 810 questions during 75 hours of recordings. In a second study, 4 boys and 3 girls were recorded weekly over a year’s duration resulting in 1,915 questions and 71 hours of recording. In both study 1 and 2, children asked a wide variety of questions during both illustrated and non-illustrated book reading. The most frequent questions (60%) centered on the book’s illustrations. Inquires about story meaning (25%) and questions about word meaning (10%) were less frequent. Questions about graphic form (e.g., letters, punctuation, printed word arrays) occurred least frequently (5%). In sum, findings substantiated that children ask many questions during storybook reading. In
the current study, the author investigates children’s ability to answer questions at a range of language levels as a measure of the child’s ability to use science language.

**Barriers to Science Instruction in Early Childhood Classrooms**

There are two barriers that limit the inclusion of science intervention in preschool classrooms. Barriers include teachers’ knowledge of science content and teachers’ understanding of developmental progression of science learning. Both barriers will be discussed in more detail in the following section.

Less than 36% of working teachers feel prepared to implement science curriculum in their classrooms (U.S. Department of Education, 2002). Also, as mentioned previously, preschool educators tend to ignore science instruction in pre-K programs (Brenneman, et al., 2009). One barrier to preschool science instruction is that early childhood educators do not always have a thorough comprehension of the science content needed to develop high-quality science activities (Darling-Hammond, 2000). Preschool educators need improved preparation and understanding of science content and training in methods to engage children in exploratory thinking in order to provide high-quality science instruction (Brenneman et al, 2009).

As an example of the need to train teachers to facilitate exploratory thinking, Krueger and Sutton (2001) studied teachers’ planning skills in correlation with student outcomes. Videotapes of classroom instruction (subjects included 50 children ages 7 to 9) were gathered from a control classroom in which teacher’s lessons were rigidly planned as well as an experimental classroom in which teachers were encouraged to be fluid in curriculum planning. Students were given pre- and post-tests. Pre-test scores were equivalent while post-test varied between groups. In the “fluid” classroom, differences
were noted in teacher behavior. Researchers observed that teachers in the “fluid” classroom used thoughtful questioning to facilitate classroom discussion. It appeared that teachers used questioning to determine a student’s thinking process and guide feedback and instruction. Researchers concluded that teachers support learning by encouraging scientific reasoning to explore concepts and explain the problem-solving process. However, many teachers are not trained to use a fluid teaching style and struggle to provide high-quality science instruction.

A second barrier to effective science instruction is that teachers must understand the developmental sequence of science and language learning. Without this knowledge, teachers may present information that is either too difficult or too easy (Shulman, 2002). In order for children to learn science and become scientifically literate, educators must choose appropriate science content and experiences to match children’s cognitive capabilities at different stages of development. Implications from research conducted by Covington and Berry (1976) report children become frustrated when there is a mismatch of content and developmental levels. Without developmentally appropriate science instruction, children may struggle to use the higher-order skills critical for science learning.

**Informational Texts as a Key Feature of Science Learning**

Adult-child shared book reading is a key feature of language learning and is a frequently occurring activity of early childhood (Dickinson & Neuman, 2006). Educational interventions in which preschool teachers use shared book readings to teach specific targeted information are effective, low cost, and able to be implemented with minimal training. For example, children whose teachers used a print referencing style
during shared book reading interactions showed statistically significant gains on print concept knowledge, alphabet knowledge, and name writing at the end of the preschool year and continued to demonstrate higher skills than the control-group children at the end of the kindergarten year (Piasta et al., in preparation).

In one study, Justice, Kaderavek, Fan, Sofka and Hunt (2009) exposed young children to alphabet concepts, letter names, and print concepts during shared book reading. Researchers implemented a 30-week book reading intervention in 23 classrooms and documented the literacy and language behaviors of 106 four-year-old children. Fourteen teachers used a print referencing (PR) style during storybook reading sessions while nine comparison teachers used a “normal” reading style (C). Results indicated that children in classrooms implementing print referencing techniques scored higher on evaluative assessments as compared to the control group; differences were observed in print organization (PR $M = 5.2$, C $M = 3.4$), print meaning (PR $M = 3.5$, C $M = 2.3$), letters (PR $M = 15.2$, C $M = 12.6$) and words (PR $M = 15.1$, C $M = 2.6$). Data demonstrated that an intervention based on shared storybook reading was an effective means to instruct young children.

Science instruction exposes children to informational texts, sometimes called “expository texts.” Expository texts differ from typical narrative books; narrative books have a story format with characters and a plot. Despite the need to expose children to variety of book types, expository texts are rarely used in preschool settings—typically less than 5%-of classroom storybook readings (Yopp & Yopp, 2006). The lack of exposure to expository texts occurs because many educators believe that children should learn to read and listen to narrative texts in the early childhood years and then progress to
reading expository texts in the later elementary grades. However, this approach can cause difficulty for students when they enter middle-elementary grades and are suddenly exposed to expository texts in science and social studies curriculum units. Experts agree that children should be exposed to expository texts from the earliest ages (Morrow & Gambrell, 2004).

Studies document the effectiveness of informational texts for promoting learning in young children. For example, Duke and Kays (1998) examined children’s knowledge of information book language. Twenty five- and six-year-old children were exposed to classroom read-alouds including both expository and narrative texts. Researchers asked children to “pretend read” an unfamiliar information book at the beginning of the school year, and again in December after 3 months of exposure to information books. Researchers reported improvements in technical vocabulary (Sept $M=.049$, Dec $M=.063$), classificatory structure (Sept $M=.008$, Dec $M=.012$), and comparative/contrastive structure (Sept $M=.000$, Dec $M=.004$). Importantly, the participants not only demonstrated increased proficiency with informational test, but the children demonstrated enjoyment during informational text reading (Duke & Kays, 1998). The results indicate the potential impact of a more intensive exposure to informational books during the preschool years.

A final important component of book reading is that shared book reading facilitates children’s vocabulary development. Specifically, Torr and Scott (2006) researched how, during shared-book reading, teachers scaffold new vocabulary and facilitate children’s interest in words and awareness of subtle differences in word meaning. In their study, a group of twenty-four children were split into two groups. In the
first group, the teachers were prompted to strictly read the book and answer questions posed by the students. In the experimental classroom, teachers were instructed to elicit a range of conversational topics during story time and ask the students questions regarding the stories and the ongoing topic of conversation. The authors reported the “enhanced reading” group engaged in more talk outside of the reading and conveyed 2102 messages in contrast to only 416 messages in the “text only” reading sessions. These data document the relationship between conversational quantity and quality and children’s participation and learning. This high-level adult-child conversational exchange has been proposed as a critical feature fostering children’s vocabulary development and literacy skills (Snow et al., 1998). This study proposes that adult-child shared book reading with an expository text can be an important component of early childhood science instruction.

**Language Features Associated with Science Learning**

Science instruction and exposure to expository texts improves children’s use and comprehension of high-level language (Varelas & Pappas, 2006). Much of this learning is supported by the kinds of questions that teachers ask during science instruction and expository book reading; this includes the adult’s use of inferential questions requiring inferencing, prediction, reasoning, and explanation (van Kleeck, 2003). Examples of high level questions include: *What do you think will happen next?* “*Why did this happen?*” “*If we put XXX and XXX together, how will XXX change?*” Expository texts are an optimal means to include high-level language into the classroom curriculum. This study uses a range of questions embedded within an expository text to document preschooler’s ability to respond to (and use) high level science language.
Science language is sometimes called inferential language; inferential language is used for analyzing, hypothesizing, and integrating ideas and information. Inferential language is used to talk about events beyond the here-and-now. Inferential language can be contrasted with literal language; literal language occurs when children are asked to discuss, describe, and/or respond to information they can readily perceive. Experts propose that high level inferential language behaviors are facilitated with exposure to science instruction (Peterson, 2009). Targeted “science language” skills include: (1) predicting/hypothesizing about subsequent events/conditions (e.g., use of causal conjunctions such as “because” and “so.”) (2) defining a word’s meaning and/or defining the purpose of an object, (3) explaining conditions that cause alternate outcomes or a solution, (4) identifying causes of occurrence/event, (5) comparing similarities/differences of objects, characters, or print, (6) formulating a generalization about events, (7) asking and answering “Wh” questions. A central premise of the current study is that teachers’ should monitor children’s use of science language, that different kinds of adult questions during book reading elicit different kinds of science language, and that children’s language is facilitated via questions eliciting science language.

An example of a study investigating the impact of adult question asking during book reading was completed by Walsh and Blewitt (2006). They examined the effects of adult questioning on children’s word acquisition during storybook readings. A group of thirty-five preschool students were each assigned to one of three conditions; (a) vocabulary eliciting questions, (b) noneliciting questions and (c) no questions (control group). Children were administered the *Peabody Picture Vocabulary Test, 3rd edition* (PPVT-III; Dunn & Dunn, 1997) and the *New Word Comprehension Test* (NWCT;
Blewitt, Rump, Shealy, & Cook, 2009) as pretest measures. Children were then to read
three different storybooks repeatedly across four reading sessions and were then tested
for production and comprehension of novel words in the final session. During the
sessions, the three groups were each asked a different variety of questions depending on
their group placement. As a post-test measure, the New Word Comprehension Test was
again administered after the sessions. Pairwise comparisons using a Tukey HSD
procedure indicated that both the “eliciting” and the “noneliciting” question conditions
yielded better performance on the NWCT than the control condition. Results also
indicated that production performance was significantly different across the three groups.
Pairwise comparisons of the means indicate that children in the “noneliciting” condition
performed better than those in the control condition, but not better than those in the
“eliciting” condition. These results support the positive impact of questioning on
children’s language performance.

Need for assessment tool to document preschoolers’ “science language”.

Currently, there are few assessments used to measure, aggregate, and compare
science learning outcomes for young children (Samarapungavan, 2009). Educators need
to have a valid and reliable method to determine which children possess adequate levels
of inferential language and which children are only able to answer concrete questions.
Once a teacher knows this information, he/she can design hands-on science units
supporting children’s higher level language development. The need to target specific
level of inferential language means that a child’s use of “science language” must be
measured periodically to determine the child’s development of inferential and abstract
language. Once science language can be measured effectively, teachers and specialists
can confidently implement interventions and monitor the progress of those in need of additional experiences and supports. To date, there is no assessment tool in place that can be used to verify children’s development of the inferential language associated with preschool science instruction.

Samarapungavan (2009) devised a study to evaluate the effectiveness of an assessment for kindergarten children. The study included 100 kindergarten children from two different schools participating in a Scientific Literacy project. Teachers and students of the experimental group engaged in a sequence of inquiry activities to explore the properties of living things and the theme of growth and development of a monarch butterfly. The experimental group was compared with a group of children in a control school; the teachers in the control school stated they did not “teach science” because of the heavy focus on literacy and numeracy. However, the teachers in the control school indicated that they sometimes included book reading on topics such as animals and seasons.

Each school was later tested using Samarapungavan’s screening assessment; the researcher demonstrated a statistically significant difference in performance scores between groups. The assessment included a 24-item objective test, broken down into two subtests that consist of the Scientific Inquiry Process, measuring young children’s functional understanding of nature and processes of scientific inquiry, and the Life Science Concepts, measuring children’s understanding of specific science concepts relating to living things and the physical world. The mean for children in experimental school was 16.91 (SD = 3.74), while the mean of the control school was 12.03 (SD = 3.07). Findings were consistent with the fact that the group receiving science learning
directly learned more than the comparison group. Findings also indicated that the assessment was sensitive to variations in science instruction. However, it should be noted that Samarapungavan’s (2009) assessment measured children’s knowledge of science content. In contrast, the proposed study argues that children’s ability to use “science language” is a separate component of science learning and should be documented along with, or prior, to children’s ability to learn science content.

The Science Learning Assessment exhibited strong reliability and validity in its design. However, the assessment was designed for kindergarten rather than preschool children. Research conducted by the National Scientific Council on the Developing Child (NSCDC, 2007) indicates that the early childhood years are critical for fostering beginning knowledge required for later educational success; the current project responds to that mandate by focusing on preschool children.

**Documenting Reliability and Validity in Child Assessments**

In order to establish that the Preschool Science Assessment Protocol (PSAP) is a sound tool, multiple measures of validity and reliability must be completed. The current study is proposed to establish preliminary evidence of validity. One essential category of validity—concurrent validity—evaluates relationships between similar variables. In this study, concurrent validity was evaluated by assessing a group of children with both the Preschool Science Assessment Protocol and the Clinical Evaluation of Language Fundamentals-Preschool 2nd edition. The CELF-P2 is a current instrument with good reliability and validity used for assessing early expressive and receptive language ability in young children. Therefore, it was hypothesized that if the PSAP was significantly
correlated with the scores of the CELF-P2, the evidence would suggest that the PSAP
displays evidence of concurrent validity.

Inter-rater reliability measures the consistency of ratings among raters by
comparing the raters’ scores with each other (Wu et al., 2007). In order to establish that
the PSAP is reliable tool, inter-rater reliability was documented by randomly choosing
twenty percent of the children’s assessments. Twenty percent of the assessments were
videotaped; videotaping permitted a second assessor to score each assessment. This
second rater independently rated the children from the video recording using the CELF-
P2 and the PSAP. Inter-rater reliability was computed by using the formula number of
(agreements/number of disagreements + agreements) X 100. Through this equation, the
researcher was able to document the correlation, with the goal of achieving a >90%
correlation. If scores on the PSAP significantly correlated with the scores of the CELF-P2
between the first and second raters’ scores, it was hypothesized that the PSAP would
display appropriate levels of inter-rater reliability.

Alternate form reliability refers to the consistency of a students’ performance on
two different versions of the same test (Popham, 2009). The current study documented
alternate form reliability by using two separate book readings (Book A and Book B). The
two assessment protocols were matched for length, mean length of utterance, and
question type. The development of the two matched protocols were designed to obtain
equivalent information for each child across the two forms. The current study was
designed to provide a preliminary source of evidence contributing to the overall validity
of the PSAP.
Implementation of the Science Assessment

The *Preschool Science Assessment Protocol* (PSAP) is a novel instrument in the field of science and language learning. The intent of the PSAP is for it to be used as an assessment tool to assess preschool children’s use of science language. The PSAP is comprised of two book-reading sessions in which a variety of questions are asked of the children including (a) causal language; (b) WH-questions; and (c) prediction questions. These areas of questioning were chosen because of the strong correlation with knowledge acquisition for school-aged children (Bloom, 1982). Having access to an authentic assessment tool, appropriate for early childhood settings, provides educators with the tools to effectively engage in the learning process implored in the education system. The PSAP will help curriculum planning as teacher’s can differentially support children at varying levels of inferential language use. The PSAP will also help identify children who are at-risk or showing delays in high-level inferential language.

Research Questions/Objectives

1. *Does the Preschool Science Assessment Protocol (PSAP) demonstrate content validity in that children’s scores on the PSAP correlate with a standardized, norm-referenced assessment of children’s language* (CELF-P2; Semel, Wiig, & Secord, 2004)?

2. *Do the two versions of the PSAP (Book 1 and Book 2) demonstrate alternate form reliability?*
Chapter Three

Methodology

In the area of science language development, contextually relevant and engaging assessment and progress monitoring tools for preschoolers are currently not available. A novel assessment instrument currently undergoing development is the Preschool Science Assessment Protocol (PSAP). The PSAP is designed to be a contextually and developmentally appropriate early science language monitoring assessment tool. The purpose of the PSAP is to document the development of children’s “science language.” The term science language, also referred to as inferential language, is used to describe the high-level language associated with predicting, describing, hypothesis generating, inferencing, describing cause-and-effect, and comparing and contrasting (see literature review above).

Research Design and Study Overview

The objective of the current study was to establish preliminary concurrent validity and alternate-form reliability for the PSAP to assist in verification of the tool for further research and use. By having access to an authentic assessment and progress-monitoring tool that is appropriate for preschoolers, educators will be able to effectively document children’s growth of science language as a result of children’s exposure to science instruction and expository texts (see literature review).

Alternate form reliability is the ability to generalize items on one test to similar items on another test. Constructing two or more similar forms of a test and administering
it to the same participants within a short period allows one to establish alternate form reliability. The administration of alternate forms should be randomized and counterbalanced. A correlation coefficient, also called the coefficient of equivalence, is calculated to compare the subject’s scores on the first form to the scores on the second. The higher the correlation coefficient, the more confident the user can be that alternate forms of the test produce equivalent results. The correlation between the alternate forms should be high for the user to infer that the content on each is equivalent (Crocker & Algina, 2006; McAfee & Leong, 2007). In the current study, alternate form reliability measures were established by using two separate book readings (Book A and Book B). Each of the books was matched for length, mean length of utterance, and question type.

In the current study, concurrent validity measures were established by assessing a group of preschool children with the *Preschool Science Assessment Protocol* and another current measure, the *Clinical Evaluation of Language Fundamentals – Preschool* (CELF-P2; Semel, Wiig, & Secord, 2004). The mean scores of the PSAP were compared to the preschool child’s performance on the subtests of the CELF-P2. In order to collect validity evidence in regards to relations to other variables, correlations were computed to measure the strength and direction of the relationship between the measures of interest. It was hypothesized that the PSAP and the CELF-P2 would correlate highly as the PSAP would document children’s use of high level language forms including question formation, use of causal conjunctions (e.g. because, so), and making inferences (e.g. prediction) and the CELF-P2 is a high quality measure of children’s language ability.

*The Clinical Evaluation of Language Fundamentals, Preschool Second Edition* (CELF-P2; Semel, Wiig, & Secord, 2004) is a current measure designed to measure used
to identify, diagnose, and perform follow-up evaluations of language deficits in children ages 3 to 6 years. The CELF-P2 was designed as a downward extension of the *Clinical Evaluation of Language Fundamentals-4th edition* (CELF-4), used for children of older ages, specifically 5-21 years of age. The CELF-P2 follows the same subtests as the CELF-4, including sentence and word structure, expressive vocabulary, directions, sentence recall, and word classes. The CELF-P2, as well as its previous versions, has demonstrated well-established measures of reliability coefficients. The CELF-4 has demonstrated a high reliability in regards to internal consistency ($r = .88$ or greater), test-retest reliability ($r = .79$ or greater across age groups and subtests), and inter-scorer reliability ($r = .97$).

Alternate-form reliability was established by determining that children’s scores on the two versions of PSAP (Book 1 = *Animal Noses*, Book 2 = *Let’s Ride Bikes*) were equivalent. The *Systematic Analysis of Language Transcription* (SALT, Miller & Chapman, 1998) was used to complete microanalyses of the children’s number of utterances and responsiveness to the examiner’s questions following language transcription of the children’s questions (see the following information on language transcription). A micro-comparison of children’s utterances in response to the two forms of the PSAP provided confirmatory data establishing alternate-form reliability.

Language transcription and the subsequent analysis of children’s utterances through language sample analysis (LSA) have enabled researchers to develop a stronger understanding of the development of language production in typical children. The process has evolved from handwritten transcription and analyses to computer software designed to improve the efficiency and accuracy of transcriptions. The *Systematic Analysis of*
Language Transcription (SALT, Miller & Chapman, 1998) is the most frequently used computer-based LSA program. The SALT program is an efficient automated analysis (Schuele, 2010). In the current study, the children’s responses were transcribed and analyzed with SALT to document the children’s total number of utterances, mean length of utterance, responses to questions, type token ratio in relation to the number of different word roots, and percentage of responses to the elicited questions.

Participants.

A total of 9 preschool students between the ages of 3 years, 6 months and 5 years, 0 months were recruited to participate in the study based on age and developmental level. Recruitment flyers and letters of consent were sent to the parents of eligible children from All God’s Children Christian Preschool in Brecksville, Ohio. The participants were recruited based on their age level; all children were reported to have typically developing language ability. Participation was voluntary based on parental consent and student identities were kept confidential by assigning each student an identification code. Test administration was completed at the child’s school during the hours of the children’s regularly scheduled preschool class for the convenience of parents as well as comfort level of the students.

Recruitment for the study was based on the following criteria. In order to be recruited for participation in the study, children met the following criteria:  a) enrolled in the preschool program, b) primary English-language speaker, c) chronological age between 3 years, 6 months and 5 years, d) parent and teacher report that the child was developing typically, and (e) report of normal hearing ability as confirmed by the child’s parents. All children for whom consent was obtained participated in the study.
Table 1

Subject's age in months

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>Age (In Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
</tr>
<tr>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>53</td>
</tr>
</tbody>
</table>

Materials: PSAP.

The Preschool Science Assessment Protocol (PSAP; Kaderavek, 2009) was comprised of 10 test items across five domains for a total of 13 possible points. The domains measured were: (a) prediction, (b) WH questioning, (c) causal relationships, (d) argumentation and (e) vocabulary.

The PSAP (Kaderavek, 2009) was administered during a natural and socially relevant activity, an adult-child shared book reading. “Animal Noses” and “Let’s Ride Bikes” are two expository books chosen for their science language based content and their preschool level focus. During the shared book reading, predetermined questions were presented to the child by the examiner, responses recorded, and a score awarded.
based on the child’s response. Scoring was varied based on the domain and task being assessed. Individual domain raw scores were calculated by summing the points in each domain. A total raw score was obtained by summing the total points across the domains.

The PSAP assessment protocol is shown in Appendix B.

In the prediction domain, there were three items addressing whether the child could predict what would happen if the story were different. Prediction items were scored on a 0-1 scale. For these items, a score of 1 represented a correct response and a 0 represented an incorrect response (i.e., child did not respond or child answered
incorrectly). The WH questioning domain consisted of two items in which the child was asked to answer a question in which they needed to apply information given in the story. The two assessment items were scored on a 0-1 scale outlined above. The child was awarded 1 point if the response was correctly related to the story and 0 points for an unrelated response or a failure to respond.

The causal relationships domain consisted of asking the child to repeat a sentence including a causal word (i.e., because). The child was awarded 1 point if the word was used in the child’s sentence and 0 points for an incorrect response or failure to respond.

In the argumentation domain, the child was asked to respond to how he or she would respond to a given false statement. The child was awarded a score of 1 if their response was correctly disputed and a 0 if the answer did not “make sense,” (i.e., was not related to the topic), if the child’s answer was not framed as an argument statement, or if the child failed to respond. The examiner scored the task in the vocabulary domain by analyzing the child’s answer to a question regarding what specific vocabulary words mean. Scores ranged from 0 (no response or entirely wrong response) to 2 (correct usage of the
definition). A score of 1 indicated that the child had some sense of the word meaning, but the answer was not completely correct.

**Data collection procedures.**

Informed consent was sought from the parents of the preschool students identified as eligible for participation in the study via an informational flyer and written consent form. These items were sent home with all eligible children for parents to review. Envelopes were provided to return the consent forms if the parents were interested in their child participating in the study. These notices were also to inform that their child’s participation, or lack of participation in the study, would not affect their child’s education in the preschool program.

The researcher, a master’s student, tested the children individually at the preschool centers. Test administration took place in two phases and was conducted over a week’s duration. The first phase was to administer the CELF-P2 and the first reading assessment of the PSAP (Book A or Book B). The second phase, conducted a few days later, consisted of another reading assessment of the PSAP (Book A or Book B). The administrations of the assessments took place individually with the evaluator. The evaluator was qualified to conduct assessments with preschool children and trained in assessment techniques specifically for young children.

During the first phase, the preschool participants were evaluated with the given assessments, the CELF-P2 and one of the PSAP book reading assessments (Book A or Book B). The total administration time for the first session was approximately 30-45 minutes per child. The next phase, another PSAP book reading assessment (Book A or Book B), took approximately 10-15 minutes per child. The two sessions were completed
within a week’s time. Presentation order of the two PSAP instruments (Book A; Book B) were randomized and counterbalanced to account for order effects as shown in Table 2.2.

Table 2

*Randomization of Testing Materials.*

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>Assessment 1</th>
<th>Assessment 2</th>
<th>Assessment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Let’s Ride Bikes</td>
<td>CELF P-2</td>
<td>Animal Noses</td>
</tr>
<tr>
<td>2</td>
<td>Animal Noses</td>
<td>CELF P-2</td>
<td>Let’s Ride Bikes</td>
</tr>
<tr>
<td>3</td>
<td>Let’s Ride Bikes</td>
<td>CELF P-2</td>
<td>Animal Noses</td>
</tr>
<tr>
<td>4</td>
<td>Let’s Ride Bikes</td>
<td>CELF P-2</td>
<td>Animal Noses</td>
</tr>
<tr>
<td>5</td>
<td>Animal Noses</td>
<td>CELF P-2</td>
<td>Let’s Ride Bikes</td>
</tr>
<tr>
<td>6</td>
<td>Let’s Ride Bikes</td>
<td>CELF P-2</td>
<td>Animal Noses</td>
</tr>
<tr>
<td>7</td>
<td>Animal Noses</td>
<td>CELF P-2</td>
<td>Let’s Ride Bikes</td>
</tr>
<tr>
<td>8</td>
<td>Let’s Ride Bikes</td>
<td>CELF P-2</td>
<td>Animal Noses</td>
</tr>
<tr>
<td>9</td>
<td>Animal Noses</td>
<td>CELF P-2</td>
<td>Let’s Ride Bikes</td>
</tr>
</tbody>
</table>

Twenty percent of the children were randomly selected for video recording. A second rater independently rated the children from the video recording using the CELF-P2 and the PSAP. Inter-rater reliability was computed by using the formula number of (agreements/number of disagreements + agreements) X 100. Inter-rater reliability on the PSAP was computed and found to be 94%; the PSAP demonstrated appropriate levels of reliability between Book A and Book B.
Chapter Four

Results

This study provided preliminary analyses of concurrent validity and alternate-form reliability for the *Preschool Science Assessment Protocol* (PSAP; Kaderavek, 2009). The goal of the study was to determine if the language variables measured by the PSAP were consistent with those on well-established, norm-referenced measure of language ability, the *Clinical Evaluation of Language Fundamentals Preschool, Second Edition* (CELF-P2; Semel, Wiig, & Secord, 2004) and to determine if the two versions of the PSAP were equivalent.

As mentioned in the section above, this study included 9 participants ages 3:8 (years:months) to 5:1, the mean age was 4:4 (SD = 5.0). Two statistical analysis protocols were completed to document (1) concurrent validity and (2) alternative-form reliability. To answer the first research question regarding concurrent validity, the researcher computed correlations between the PSAP and CELF-P2 using partial correlation coefficient analysis. Partial correlation analysis reports the correlation between two variables after removing the effects of other variables—in this case children’s age. Partial correlation computation identifies correlations masked by the effect of other variables. It was hypothesized that the PSAP would highly correlate with the CELF-P2 once the factor of a child’s age was statistically controlled via the partial correlation process.
In order to examine the concurrent validity of the PSAP, the researcher computed the partial correlation between children’s scores resulting from the PSAP and the CELF-P2. As previously described, concurrent validity compares scores on one measure (the PSAP) with current performance on other independently established measure (the CELF-P2) resulting in a correlation coefficient (Cizek, Rosenberg, & Koons, 2008; Ellis & Blustein, 1991; Sattler, 2001). Data analysis was conducted using the Statistical Package for the Social Sciences (SPSS) software to establish measures of concurrent validity.

Concurrent validity is typically reported as a correlation coefficient. In this study, partial correlation statistics were employed to control for the issue of children’s age, to calculate a correlation coefficient, to determine if a linear relationship existed, and to describe the strength of the relationship between the variables. Correlation coefficients can range from +1.00 to -1.00. The closer the correlation is to 1 or -1, the stronger the relationship. Positive coefficients indicate positive relationships (i.e. as scores on one test increase, scores on the other test increase) while negative coefficients indicate inverse relationships (i.e. as scores on one test increase, scores on the other decrease). A correlation does not imply that one variable causes the other, simply that there is a relationship between the two (Irwin et. al, 2008).

The partial correlation computation resulted in a significant correlation for both Book 1: *Noses at work* (*pr = .81, p = .01*) and Book 2: *Riding Bikes* (*pr = .71, p = .05*). A correlation of >.7 is considered a strong correlation (Schiavetti & Metz, 2002); both forms of the PSAP were strongly correlated with the results of the CELF-P2.

In response to the second research question (i.e., alternate-form reliability), three sub-analyses were completed. First, the researcher examined the alternate form reliability
by computing the correlation between Book 1 and Book 2. Second, as a different way to explore alternative form reliability, the researcher compared the children’s mean scores between the two versions of the PSAP versions (Book 1 and Book 2) to see if there was a significant difference in performance between the two books. Third, to complete a fine-grained analysis of children’s responses during Book 1 and Book 2, the researcher compared the children’s mean length of utterance and number of responses to the examiner’s questions using the Systematic Analysis of Language Transcription (SALT, Miller & Chapman, 1991). For the second and third set of analyses, a nonparametric test, the Wilcoxon Signed-Ranks Test, was used to consider the difference between children’s performance. Nonparametric analyses are more appropriate with small sample sizes that may lack a normal distribution of scores (Schiavetti & Metz, 2002).

The partial correlation coefficient (controlling for children’s age) revealed a high significant correlation ($r = .742, p = .035$) when comparing the total scores for Book 1 and Book 2. Further, when submitted to a Wilcoxon Signed Rank Test, comparison of children’s mean scores on the PSAP documented that the children’s performances on the two versions were not significantly different ($T = .94, p = .367$). Finally, fine-grained analyses demonstrated that (a) children’s MLU across the two PSAP forms were not significantly different (Book 1 $M = 4.6, SD = 1.7$; Book 2 $M = 5.3, SD = 1.2$; $T = 1.2, p = 2.6$) and (b) children’s total number of responses to the examiner’s questions were not significantly different (Book 1 $M = 12.2, SD = 2.3$; Book 2 $M = 13.7, SD = 4.1$; $T = .95, p = .37$).

As discussed in the previous chapter of this paper, a child’s mean length of utterance (MLU) is considered one of the most robust indices of young children’s
language acquisition (Rice et al., 2010). MLU refers to the number of words of morphemes in a child’s spontaneous utterances. In the current study, data reflecting the (a) child’s MLU and (b) the child’s total number of responses were computed via SALT to determine if a child’s responsiveness was (or was not) different significantly between the two PSAP forms. It was hypothesized that children’s responsiveness would not be significantly different between the two forms, a finding that helps establish the PSAP’s alternate-form reliability. The subjects’ individual scores on the CELF-P2, PSAP (Book 1), and PSAP (Book 2) are shown in Table 3. The mean raw score on the PSAP (Book 1) for the total sample was 5.67 (n = 9, SD = 2.2); the mean raw score for the PSAP (Book 2) was 4.8 (n = 9, SD = 1.8). The mean standard score on the CELF-P2 was 105.7 (n = 9, SD = 13.2).
Table 3

*Individual scores of testing materials*

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>CELF P-2 Standard Score</th>
<th>Book 1 Animal Noses</th>
<th>Book 1 Let's Ride Bikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>112</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>106</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>129</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>98</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>123</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>94</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>
Chapter Five

Discussion

The purpose of this study was twofold. First, the study examined the concurrent validity of the *Preschool Science Assessment Protocol* (PSAP; Kaderavek, 2009) by comparing children’s performance on the PSAP with the *Clinical Evaluation of Language Fundamentals – Preschool, Second Edition* (CELF-P2; Semel, Wiig, & Secord, 2004). Second, the study examined the alternate form reliability of the PSAP to determine the equivalency of the two forms of the assessment.

In response to research question one, “*Does the PSAP demonstrate content validity in that children’s scores on the PSAP will correlate with a standardized, norm-referenced assessment of children’s language?*” findings demonstrated that the PSAP highly correlated with the results of the CELF-P2. The subtest scores on the CELF-P2 are designed to measure children’s sentence and word structure, expressive vocabulary, directions, sentence recall, and word classes. The PSAP correlated most strongly with the composite score of the CELF-P2, supporting that the PSAP appears to be an effective measure of language skills in preschoolers.

Results regarding research question two; “*Do the two versions of the PSAP (Book 1 and Book 2) demonstrate alternate form reliability?*” indicated that the children’s performances on the two versions were not significantly different ($T = .94, p = .367$). Further, children’s overall language performance did not differ between the two forms of the PSAP. Specifically, there was not a significant difference in the children’s MLU.
(Book 1 $M = 4.6$, $SD = 1.7$; Book 2 $M = 5.3$, $SD = 1.2$; $T = 1.2$, $p = 2.6$) or total number of responses to the examiner’s questions (Book 1 $M = 12.2$, $SD = 2.3$; Book 2 $M = 13.7$, $SD = 4.1$; $T = .95$, $p = .37$). The equivalence of the children’s responses across the forms is an additional verification of the tool’s reliability.

Results of the present study support the utility of the PSAP as an assessment tool for measuring early science language in preschool students. An examination of validity evidence indicated that the PSAP has high levels of concurrent validity. These results provide support for the valid use of the PSAP as an assessment tool for preschoolers. The results are encouraging, as there is a significant need for an authentic early science language-monitoring tool that can be utilized in preschool settings to assist in the prevention and monitoring of science language learning (Nichols, & Berliner, 2005).

Due to the child-friendly structure of the PSAP, children responded favorably to the assessment. The majority of the children were comfortable and familiar with book reading activities and appeared confident during the assessment. When requested to answer specific questions in the assessment, the children readily complied and often shared unrelated information regarding their personal experiences at home or in the classroom. In addition, while reading the PSAP short story “Let’s Ride Bikes”, many of the children related to the story, offering information about their own experiences (i.e., “I have a blue bike and my sister’s is pink.” or “My dad lets me ride my bike alone as long as I have a helmet on.”). Children’s natural reaction to the story provided observational evidence that the PSAP was perceived as a meaningful and relevant task.

The results from other’s validity studies provide credence to the findings in this study. For example, Justice, Bowles, Pence, and Gosse (2010), evaluated the validity of
an early assessment of children’s language abilities, the *Narrative Assessment Protocol* (NAP). Justice and her colleagues examined the NAP’s concurrent validity by correlating children’s responses on the NAP with children’s performance on the CELF-P2. In this particular study, scores on both forms were consistently and moderately correlated with the CELF-P2 scores. Specifically, correlations for the NAP sum score and the Core Language Composite were medium in size, $r = .35, p < .001$ and $r = .34, p < .001$ respectively. Given the results of the correlations, Justice, Bowles, Pence, and Goose (2010) concluded that the NAP demonstrated concurrent validity with another standard language measure (CELF-P2). It should be noted that the correlations in the current study were higher than those reported with the NAP (i.e., in the current study the correlations were $r .71$ and $81$). As such, the interpretation of the PSAP’s validity is supported by other current work in preschool assessment.

In an example of reliability documentation, Cocker and Ritchey (2010) evaluated writing abilities in preschoolers using a curriculum-based measurement (CBM). Cocker and Ritchey examined the psychometric properties and utility of the CBM by administering the assessment in various forms to preschool aged children. In order to examine alternate-form reliability, the researchers administered different probes of the CBM. Subsequently, they documented alternative form reliability by calculating Pearson correlation coefficient between various probes. The authors reported coefficients ranging between .65 and .81 with most of their correlations in the .70 range; they interpreted their levels as documenting satisfactory reliability for the CBM tool. In the current study, the author computed correlation between forms at .742. Given that other test developers are
using correlation results of similar magnitude, the reliability results of the PSAP are promising.

**Limitations and Future Research**

As in all research studies, some limitations affected the results found in the current study. The sample size of this study is one limitation. Only one small preschool was used and the study examined the performance of 9 participants. This is a small sample in comparison to the total number of students in a typical preschool program. Second, the children in this study may not represent a typical preschool population. The preschool was a Christian preschool and the students were from families who are middle to high social-economic status. Results of this study may not be generalized to the broader preschool population. Using this tool with a larger, more representational sample is a goal for future research with the PSAP.

Another limitation of the current study was the specific scheduling of the assessment times and the requirement that children needed to leave the classroom to complete the assessment task. It is possible that some children may have been anxious about leaving the ongoing activity and may not have given their “best performance.” In the future, it may be advisable for assessments to be completed as soon as the children entered the classroom in the morning to minimize this effect.

Although a longitudinal (i.e., “predictive”) component was beyond the scope of this project, it should be emphasized that the current study did not examine the diagnostic properties of the children’s use of science language. That is to say, it did not determine if children who perform less well on the PSAP perform less well on science activities in the classroom or if preschoolers who score poorly on the PSAP are at risk for science
learning in later elementary school. The ability to identify children who are at risk for science learning and the ability to document children’s development as a result of high-quality science instruction is the ultimate goal for the PSAP. This is an avenue for future research.

**Final Summary and Conclusions**

Recent research has demonstrated that exposure to science instruction in the preschool years in turn improves children’s literacy development and facilitates language development (Andersen et al., 2009). Research also indicates the strong relationship between preschool science instruction and language development. Given recent reports regarding the U.S ranking of science knowledge compared to other children in developing countries (PISA, 2006) as well as the current education initiatives put in place by President Obama (Obama, Speeches and Remarks, November 29, 2009) the significant need for an authentic early science language-monitoring tool has grown. This study provided preliminary evidence that the *Preschool Science Assessment Protocol* is a viable tool for assessing young children’s science language.

The purpose of the PSAP is to document the development of children’s “science language” while providing information of five aspects of children’s inferential learning abilities including prediction, WH questioning, causal relationships, argumentation, and vocabulary skills. Through the development of the assessment, early childhood educators gain the ability to utilize this tool to gain information regarding which children posses adequate levels of inferential language in order to become a “science learner.” Results of the current study established measures of content validity and alternate-form reliability.
for the PSAP; suggesting it is an appropriate language learning assessment tool in which to use in assessing young children.

The PSAP has the potential of assisting early childhood educators to engage in screenings of students in order to identify those who are at-risk or not developing age-appropriate science language learning skills. The PSAP also could assist educators to effectively document children’s growth of science language as a result of their exposure to science instruction and expository texts. In summary, findings of the present pilot study support the utility of the PSAP as an assessment tool for measuring early science language in preschool students. An examination of concurrent validity and alternate form reliability indicate a high level of correlation. The results of this study are encouraging in the development of an authentic early science language monitoring tool, specifically the development of the PSAP.
References


Appendix A

Permission Forms

Dear Parents,

Thank you for allowing your child to participate in the preschool science study being conducted by the University of Toledo, Department of Early Childhood, Physical, and Special Education. Your child will be participating in several activities over the next several weeks involving storybook reading and other related tasks. In order to complete the data collection process, I am requesting that you fill out the following questionnaire and return it to school with your child. All information will be kept confidential. Thank you for your assistance.

Sincerely,

[Signature]

John N. Apdnen, Professor
University of Toledo
INFORMED CONSENT FORM
Preschool Science Book Reading

Principal Investigator: Joan N. Kadaverk, Professor
Email: Joan.Kadaverk@utoledo.edu

Purpose: Your child is invited to participate in a research project entitled, Preschool Science Book Reading, a study being conducted by the University of Toledo under the direction of Dr. Joan Kadaverk. The purpose of this study is to better understand how preschool children respond to informational books (e.g., books about simple science concepts).

Description of Procedures: This research will take place at your child’s preschool. First, you will be asked to complete some basic information (e.g., your child’s age, birthdate, etc.). Then, if you grant permission for your child to participate in the study, he or she will interact with a trained researcher who will (a) administer a language assessment (lasting 30 minutes) and (b) read two simple preschool level “science books” and talk to your child about the information in the books. The books will be on topics such as animals or plants. The interactions will take place on two separate days to make sure your child is not bored or tired. Overall, your child’s participation will take about 50 minutes and take place over a two-week period. Your child’s conversation about the science books will be audio taped so that his or her responses can be verified by another member of the research team. Following this accuracy check, the audiotape will be destroyed.

Permission to record: Will you permit the researcher to audio record during this research procedure?

YES NO

Initial Here

Potential Risks: There are minimal risks to participation in this study, including loss of confidentiality. It is unlikely, but your child may feel stressed or uncomfortable during the assessment. If this occurs, he or she will not be expected to continue. Typically children enjoy interacting with the adult researcher and are eager to share what they know; the activities are designed to be engaging and interesting to young preschool children (e.g., looking at pictures, talking about the story). The researchers work with the classroom teacher to minimize any disruption to the daily classroom schedule.

Potential Benefits: The benefit to your child is that he or she has the opportunity to interact with an interested and pleasant adult. The information from this study has an educational benefit by helping researchers understand more about how children react and talk about information texts.

Confidentiality: The researchers will make every effort to prevent anyone who is not on the research team from knowing that you provided this information, or what that information is. The consent forms with signatures will be kept separate from responses, the final information logs will not include names and will be presented to others only when combined with other children’s responses. Your child’s information will be coded with a number and the number (not his/her name) will be used to record the data. Your child’s name or school name will not be used or released.

Child’s Information Requested

Date: 11/5/07

47
Voluntary Participation: Your refusal to participate in this study will involve no penalty or loss of benefits to which you are otherwise entitled and will not affect your relationship with The University of Toledo or your child's preschool program. In addition, you may ask for your child to discontinue participation at any time without any penalty or loss of benefits.

Contact Information: Before you decide to accept this invitation to take part in this study, you may ask any questions that you might have. If you have any questions at any time before, during or after your child's participation you should contact Dr. Joan Kaderavek, Professor, 419 630 2506, email: Joan.Kaderavek@utoledo.edu

If you have questions beyond those answered by the research team or your rights as a research subject or research-related injuries, please feel free to contact the Chairperson of the SBE Institutional Review Board, Dr. Barbara Chesney, in the Office of Research on the main campus at (419) 530-3844.

Before you sign this form, please ask any questions on any aspect of this study that is unclear to you. You may take as much time as necessary to think it over.

SIGNATURE SECTION — Please read carefully

You are making a decision whether or not to your child should participate in this research study. Your signature indicates that you have read the information provided above, you have had all your questions answered, and you have decided to allow your child to take part in this research.

The date you sign this document to enroll in this study, that is, today's date must fall between the dates indicated at the bottom of the page.

Child's Name (please print)  Parent or Guardian Name (Please Print)  Date

Parent or Guardian Signature  Person's Signature Who Obtained Consent (Member of Research Team)  Date

THE UNIVERSITY OF TOLEDO
SOCIAL, BEHAVIORAL & EDUCATIONAL INSTITUTIONAL REVIEW BOARD

The research project described in this consent form and the form itself have been reviewed and approved by the University of Toledo Social, Behavioral & Educational Review Board (SBE IRB) for the period of time specified below.

SBE IRB #: 104785  Approved Number of Subjects: 15
Project Start Date: 02/05/10  Project Expiration Date: 02/05/11

Barbara Chesney, Ph.D., Chair
UT Social Behavioral & Educational IRB
STUDENT INFORMATION:

Child's Name: ____________________________  Parents Names: ____________________________

Gender: □ Male  □ Female  Date of Birth: ____________________________

Address: ____________________________  City/State/Zip: ____________________________

Phone: ____________________________  Email: ____________________________

Ethnicity/Race:  □ White  □ Hispanic  □ Black  □ Am Indian  □ Asian  □ Multi-Racial  □ Other

Child's Native Language (if not English): ____________________________

Parent’s Native Language (if different than child): ____________________________

Languages other than English spoken in the home: ____________________________

Mother’s Educational Level: ____________________________  Father’s Educational Level: ____________________________

□ Did not graduate high school  □ Did not graduate high school
□ High school  □ High school
□ Some college  □ Some college
□ College graduate  □ College graduate
□ Graduate school  □ Graduate school

Mother’s Occupation: ____________________________  Father’s Occupation: ____________________________

Sibling living with the child:

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Relationship (i.e. brother/sister)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Does your child have a history of speech/language, hearing, or developmental delays?

□ Yes  □ No

If Yes, please describe: ____________________________________________________________
Appendix B  
PSAP Form: *Noses at Work*

<table>
<thead>
<tr>
<th>SCORING: WH-QUESTION: 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION TWO TIMES)</th>
<th>WH</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAY</td>
<td>What if I said: Animals don’t need air! What would you say back to me? [IF CHILD SAYS SOMETHING, THEN SAY: “WHY DO YOU THINK THAT?”]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCORING: ARGUMENTATION (DON’T SCORE THIS YET) (CAN REPEAT QUESTION [BOTH PARTS] TWO TIMES)</td>
<td>ARG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>READ</td>
<td>Hippos stick their wide noses above water. They take a big breath before sinking back under.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASK</td>
<td>Say this sentence: “The hippo takes a big breath because he’s going underwater”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCORING: CAUSAL: 1 point if correct, 0 if incorrect. (CAN REPEAT DIRECTIONS 2X)</td>
<td>CAUSAL</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>READ</td>
<td>Sharks smell food underwater with their noses. Sharks have noses on their snouts.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>READ</td>
<td>Elephants lift their long trunks high in the air. They smell where to find water holes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAY</td>
<td><em>Why do elephants lift their trunks?</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCORING: WH-QUESTION: 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION 2X)</td>
<td>WH 0 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>READ</td>
<td>Pigs sniff the ground to find food. Then they use their noses to dig for the food.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAY</td>
<td><em>What do you think will happen when the pig smells some food?</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCORING: PREDICTION: 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION 2X)</td>
<td>PRE 0 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>READ</td>
<td>Awesome animal noses. Big or small, long or short, noses help animals breathe and smell.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAY</td>
<td><em>Say this sentence: Animals have noses so they can breathe and smell.</em></td>
<td></td>
</tr>
</tbody>
</table>
**SCORING**: CAUSAL: 1 point if correct, 0 if incorrect. (CAN REPEAT DIRECTIONS 2X)  

<table>
<thead>
<tr>
<th>Pre</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
</table>

**CHILD SAY (WRITE IN AFTERWARDS FROM TAPE):**

**QUESTIONS:**

<table>
<thead>
<tr>
<th></th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What does SNEAK mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)</td>
</tr>
</tbody>
</table>

Child says: (Write in later)

<table>
<thead>
<tr>
<th></th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>What does NOSTRIL mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)</td>
</tr>
</tbody>
</table>

Child says: (Write in later)

<table>
<thead>
<tr>
<th></th>
<th>SCORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What does SCENT mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)</td>
</tr>
</tbody>
</table>

Child says: (Write in later)

**TOTAL PREDICTION =**

**TOTAL WH =**

**TOTAL CAUSAL =**

**ARGUEMENATION =**

**VOCABULARY TOTAL =**
Appendix C

PSAP Form: *Let’s Ride Bikes*

<table>
<thead>
<tr>
<th>COVER</th>
<th>SAY:</th>
<th>Let’s read a book about bikes!</th>
</tr>
</thead>
<tbody>
<tr>
<td>READ</td>
<td>Let’s Ride Bikes</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>READ</td>
<td>Riding bikes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pedal fast! Pedal slow. It is fun to ride bikes with friends.</td>
</tr>
<tr>
<td>2</td>
<td>READ</td>
<td>Bikes roll best on flat, smooth ground. Paved paths are easy places to ride.</td>
</tr>
<tr>
<td></td>
<td>ASK</td>
<td><em>Say this sentence: “Bikers like to ride on sidewalks because they are smooth”</em></td>
</tr>
<tr>
<td></td>
<td>SCORING</td>
<td>CAUSAL: 1 point if correct, 0 if incorrect. (CAN REPEAT DIRECTIONS 2X)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAUSAL</td>
</tr>
<tr>
<td></td>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>READ</td>
<td>Some bikes race on dirt and grass. Knobby tires roll fast on the bumpy ground.</td>
</tr>
<tr>
<td>READ</td>
<td>Training wheels help new riders learn to balance. Riding a bike takes lots of practice.</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>SAY</td>
<td>What do you think might happen if the little boy doesn’t use training wheels?</td>
<td></td>
</tr>
<tr>
<td>SCORING</td>
<td>PREDICTION QUESTION: 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION TWO TIMES)</td>
<td></td>
</tr>
<tr>
<td>PRE</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>Read</td>
<td>Say</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>6</td>
<td>Biking safety.</td>
<td>Why do bike riders use hand signals?</td>
</tr>
<tr>
<td>7</td>
<td>Safe bike riders are easy to see. They wear bright colors and use reflectors.</td>
<td>Why do the children wear bright colors and reflectors?</td>
</tr>
<tr>
<td>8</td>
<td>Helmets cover riders’ heads. The hard shell protects them in crashes or falls.</td>
<td>What do you think might happen if the bike rider falls and he does not have a helmet on?</td>
</tr>
</tbody>
</table>

**SCORING**

- **WH-QUESTION:** 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION TWO TIMES)
- **PREDICTION:** 1 point if correct, 0 if incorrect. (CAN REPEAT QUESTION 2X)

**CHILD SAYS (WRITE IN AFTERWARDS FROM TAPE):**
Having fun.

Old or young, biking is fun for everyone. Let’s ride bikes!

QUESTIONS:

1. What does BALANCE mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)

Child says: (Write in later)

2. What does PAVED mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)

Child says: (Write in later)

1. What does REFLECTOR mean? (SCORING: LEAVE BLANK FOR NOW) (CAN REPEAT 2X)

Child says: (Write in later)

TOTAL PREDICTION =
TOTAL WH =
TOTAL CAUSAL =
ARGUemenATION =
VOCABULARY TOTAL =
Appendix D

Internal Revue Board Forms

<table>
<thead>
<tr>
<th>A. STUDY INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of Submission:</td>
</tr>
<tr>
<td>Study Title:</td>
</tr>
<tr>
<td>Principal Investigator: Joe ReGreen</td>
</tr>
<tr>
<td>Department:</td>
</tr>
<tr>
<td>Phone #:</td>
</tr>
<tr>
<td>Student Investigator:</td>
</tr>
<tr>
<td>Student's Phone #:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. STUDY FUNDING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding Source:</td>
</tr>
<tr>
<td>Incident number:</td>
</tr>
</tbody>
</table>

*If not university funded, please complete the following*

<table>
<thead>
<tr>
<th>Agency/Company Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agency/Company Address:</td>
</tr>
<tr>
<td>Grant title if different from the proposed title</td>
</tr>
<tr>
<td>Grant account #:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. PERFORMANCE SITE(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Site Name:</td>
</tr>
<tr>
<td>Address:</td>
</tr>
</tbody>
</table>

PSAP Proposal Application Page 1 of 11 Revised 10.30.09

57
### D. STUDY PERSONNEL

Please list all study personnel involved in the conduct of this study. Anyone who is "engaged in research" must be listed below. This includes all study personnel who interact or intervene with subjects and/or have access to subjects' identifiable private information. This list may be different (often longer) than the key personnel list included in the grant.

Check the box under UT for each person who is affiliated with UT. Only UT faculty, staff, students, or registered volunteers are considered "UT-affiliated" and thus covered by the UT IRB review. All non-affiliated study personnel must have their appointment reviewed by the appropriate IRB and cannot work on the project until the IRB approval is on file.

All study personnel must complete the required training in human subject research and provide the IRB office with the original training form and certificates of completion. The IRB will not review a study without the forms on file for all research personnel. The training forms can be found at [http://mdco.edu/officeofresearch/IRBtraining.htm](http://mdco.edu/officeofresearch/IRBtraining.htm)

<table>
<thead>
<tr>
<th>Name</th>
<th>UT</th>
<th>REGIST ID</th>
<th>PARTICIPANT'S TITLE IN THIS RESEARCH</th>
<th>PARTICIPANT'S ROLE IN THIS RESEARCH</th>
<th>TRAINING COMPLETED</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joan Rodriguez</td>
<td>☑</td>
<td>060621FF1</td>
<td>Principal investigator</td>
<td>Direct project, assisted as needed to all other listed non-principal investigators</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andrew Martinez</td>
<td>☑</td>
<td>078499BI2</td>
<td>Master's student, research assistant</td>
<td>Review subject, data entry, answer phone calls, statistical analysis, narrative data</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*Example of required information, please place above example when filling in table)*

### E. CONFLICT OF INTEREST

All Study Personnel listed in Section D must submit a completed Financial Conflict of Interest (FCO) Form.

The Financial COI form can be found at [http://mdco.edu/research/FCIForm2012.doc](http://mdco.edu/research/FCIForm2012.doc)

If there is any real or apparent conflict of interest on the part of any study personnel (e.g., stock or stock options, interest in technology, consultant to sponsor)? Yes ☑ No ☑ If yes, please explain.

### F. METHODS AND PROCEDURES (Please read carefully)

This section must be written in layman’s terms so that it can be understood by all members of the IRB.

Include sufficient detail so that the scope of your project can be understood but not so extensive as would be expected in a grant proposal or journal article.
1. Describe briefly the background and significance of the study. Significance of this research: Recent reports state that U.S. adolescents are ranked 25th in science knowledge. U.S. students from countries like Croatia, the Czech Republic, and Liechtenstein are ahead of us in developmentally advanced countries (PRSA, 2008). More recently, President Obama—in his speech on "Education to Innovate"—indicates that federal dollars will be used to bring science and math education to all U.S. children— even preschoolers (November 29, 2009). This project will develop and pilot an assessment tool to verify preschoolers' use of "science language." The Need for Science Learning in Early Childhood: Despite the need to increase science skills of U.S. children and the documented need to begin science education at the earliest ages, preschool educators tend to ignore science instruction in pre-K programs (Bremser et al., 2006). Science activities expand children's higher order cognitive and language abilities by requiring children to plan, describe, predict, make hypotheses, and explore principles of cause-and-effect. Science instruction also creates opportunities for young children to exchange information, ask and answer questions, need for information, make charts and graphs, and make observations and dictate results (French, 2004). Research demonstrates that exposure to science instruction improves children's literacy development, facilitates language development, and can accelerate overall academic progress (Anderson et al., 2009).

Informational Texts as a Key Feature of Science Learning: Adult-child shared book readings are a key means to foster child learning (Dickinson & Neuman, 2008). Educational interventions using children's storybooks are effective and low cost. For example, Justice et al. (2003) implemented a 30-week book reading intervention in a randomized, control group design experiment. The goal was to expose young children to alphabet concepts, letter names, and so forth. Children whose teachers used the trained book reading strategy showed statistically significant gains on the targeted skills and continued to demonstrate higher skills than the control group children at the end of the kindergarten year. These data demonstrated that an intervention focused on shared storybook reading is an effective means to instruct young children.

Science instruction exposes children to informational texts, sometimes called "expository texts." Expository texts differ from typical narrative books, narrative books have a story format with characters and a plot. Despite the need to expose children to a variety of book types, expository texts are rarely used in preschool settings—typically less than 5% of classroom storybook readings (Yopp & Yopp, 2005). The lack of exposure to expository texts causes difficulty for students when they enter middle-elementary grades and are suddenly exposed to expository texts. Experts agree that children should be exposed to expository texts from the earliest ages (Morrow & Gambrell, 2004). Science instruction typically includes the use of expository texts; thus the high-level language demonstrated during shared expository book reading gauges children's proficiency in the use of "science language."

Language Features Associated with Science Learning: Science instruction and exposure to expository texts improves children's use and comprehension of high-level language, sometimes referred to as "science language" or also "inference language" (Yopp & Yopp, 2005). Inference language is supported when teachers ask questions requiring prediction, reasoning, and explanation (Van Leeuwen, 2003). Examples of high-level questions include: What do you think will happen next? "Why did this happen?" "If we put XXX and YYY together, how will ZZZ change?" Inferential language is used for analyzing, hypothesizing, and integrating ideas and information. Experts propose that high-level inferential language behaviors are facilitated with exposure to science instruction (Pellegrino, 2000). Important "science language" skills include: (1) predicting/hypothesizing about subsequent events/conditions (e.g., use of causal constructions such as "because" and "so"), (2) defining a word's meaning and/or defining the purpose of an object, (3) explaining conditions that cause alternate outcomes or a solution, (4) identifying causes of occurrence/event, (5) comparing similarities and differences of objects, characters, or print, (6) formulating a generalization about events, (7) asking and answering "Why" questions.

Aims of the Assessment Tool: Document Preschoolers’ "Science Language." Due to the national focus on early science education, there is a growing need for the development of early childhood education science curriculum. However, to date there are no assessment tools to verify children's development of inferential language associated with science instruction. This project will lead to the development of an assessment tool that can then be used to document children's development of "science language."

2. What is the objective of the study?

a) To pilot the Preschool Science Assessment Protocol (PSAP), an assessment tool designed to assess preschool children’s use of science language during an expository book reading and (b) To document the reliability and validity of the assessment tool.

3. Describe the study design and all procedures (sequentially) to which human subjects will be exposed.

Following approval by the IRB Human Subjects Review Board, the research team will conduct two evaluations of the PSAP including: (a) test 15 children with two expository book protocols for test-retest reliability of equivalent forms administered across a 2-week period, randomized and counterbalanced presentations using correlation coefficients (Sheahean, Matoros, & Maramon, 2002) to determine the concurrent validity of the PSAP. Concurrent validity assesses whether the skill measured by one test is consistent with another assessment tool. We will compare the children's mean performance on the PSAP with their performance on the Child Evaluation of Preschool Language—Preschool (CELP-P; Wiig, Secord, & Semel, 2004). The CELP-P is a well-established valid, reliable measure of receptive and expressive language ability with internal consistency reliability coefficients exceeding .77, and test-retest coefficients ranging from .75 to .94.
For the current study, the children will first be presented with the CELF-P, the CELF-P is a norm-referenced, frequently used, language assessment designed to be used with preschool children. The CELF-P will take approximately 90 minutes to administer. Following the CELF-P administration the preschooler will interact with the assessor with the Preschool Science Assessment Protocol (PSAP). The PSAP will include (reading the child an informational book written for preschool children e.g. “Let’s talk about animal noses”). During the book reading the assessor will ask the child questions designed to gauge the child’s use of “science language” (e.g. inferential language). (See example questions below). The PSAP will take approximately 5-10 minutes to administer. Do another toy (within a two week period) a second PSAP book will be read. The assessors are (a) a faculty research or (b) a graduate student in speech-language pathology. Both individuals have received training in administering assessment protocols to young children. The assessment protocols are specifically designed to be engaging and interesting to young preschool children.

**EXAMPLE OF THE PSAP PROTOCOL USING THE BOOK: “LET’S TALK ABOUT ANIMAL NOSES”**

<table>
<thead>
<tr>
<th>TEXT</th>
<th>ADULT QUESTION</th>
<th>INFERENTIAL LANGUAGE ASSESED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animals breathe and smell with their noses.</td>
<td>The rabbit sneaks closer.</td>
<td>What does “sneak” mean? Assesses child’s ability to infer vocabulary meaning from exposure to expository text</td>
</tr>
<tr>
<td>Scents tell animals what’s going on around them. Sniff, sniff.</td>
<td>The rabbit’s nose wiggles.</td>
<td>Why do you think might happen if the rabbit doesn’t run away? Assesses child’s ability to predict</td>
</tr>
<tr>
<td>A fox smells a rabbit.</td>
<td>It smells the fox.</td>
<td>Why do horses take in a lot of air? Assesses child’s ability to respond to Wh-question</td>
</tr>
<tr>
<td></td>
<td>Then it hops away to safety.</td>
<td>Say this sentence, “The hippo takes a big breath because he’s going under water.” Assesses child’s ability to repeat causal conjunction “BECAUSE”</td>
</tr>
<tr>
<td></td>
<td>Horses take in a lot of air with their large nostrils.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The extra air helps them run faster and farther.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hippos stick their wide noses above water.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>They take a big breath before sinking back under.</td>
<td></td>
</tr>
</tbody>
</table>

G. SURVEYS AND QUESTIONNAIRES

☐ Surveys/Questionnaire (go to A) ☐ Racial Database (go to B) ☐ Other, Briefly explain:

A. Surveys and Questionnaires. You must attach a copy of each survey or questionnaire.

1. What type of survey or questionnaire will be used?

   Demographic Questionnaire, (see attached)
2. Describe the setting and mode of administration for the instrument (e.g., by phone, one-on-one, group) and the procedures for maintaining privacy and confidentiality (e.g., anonymity). Include duration, intervals of administration, and overall length of participation.

The questionnaires will be sent home in the children’s book bags. The demographic data questionnaires will be coded by randomly assigned student numbers to ensure confidentiality and will be kept in a locked cabinet in the principal investigator’s office.

| B. Records or Data Review. This includes existing material such as archival records, databases etc. |
|---|---|
| 1. What kinds of records will you review? What is the source of the records? |
| NONE |
| 2. Will you have contact or interaction with the subjects from whom the data are collected? |
| N/A |
| 3. Will you be recording identifiers (information that could potentially identify human subjects)? |
| N/A |
| 4. What is the timeframe of the records that you plan to review? Please define the time span. (Examples: from 2/1/2007 – 2/1/2008) |
| N/A |

H. RISK/BENEFIT ASSESSMENT

1. Describe in detail any potential risks/adverse events associated with each intervention or research procedure using the table below. Assess the degree of risk and likelihood of such risk (low, moderate, high).

<table>
<thead>
<tr>
<th>Potential Risks/Adverse Events (psychological, social, economic, legal)</th>
<th>Degree of Risk (minimal, more than minimal)</th>
<th>Likelihood of Adverse Events (low, moderate, high)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children could be stressed by assessments</td>
<td>Minimal</td>
<td>Low (evaluator is trained to work with young children, all tasks are designed to be used with young children and will be completed in a familiar setting).</td>
</tr>
</tbody>
</table>

3. What is the investigator’s overall assessment of the risk classification of the study? (none, minimal, or more than minimal risk)? According to DHHS/FDA regulations, minimal risk means the “probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.”

None/minimal. The tasks are typical of early childhood classrooms and classrooms (e.g., raising pictures, pointing to pictures, responding to adult questions, etc.). The evaluator is qualified to conduct preschool assessment; the children will all be typically developing preschoolers and the tasks are commonly occurring and familiar to usual children. Children who do not want to participate are any given day, will not be required to continue with the assessment tasks.

3. What procedures will be utilized to prevent/minimize any potential risks or discomfort (physical, social, psychological, economic, and legal)?

Children will be able to complete the assessments over several sessions if they are bored, stressed or upset in any way. They will be tested in familiar school settings and not singled out from their peers. Teachers may accompany the child if requested by the child to maximize comfort level. Teachers will be told that they should only send the participation form home with children that they believe will be comfortable with the assessment protocol.
4. Describe the potential direct benefits subjects may receive as a result of their participation.

Children spend one-on-one time with a warm and attentive adult. Children are exposed to high-quality adult-child interactions and will be engaging in the solution typical of early childhood settings. Children will spend time reading a book with an adult.

5. Describe the potential benefits to society that may be expected from this research. Social benefits generally refer to the advancement of scientific knowledge, and/or possible benefit(s) to future subjects.

There is a strong need for more appropriate assessment tools to assess young children’s learning of “science language” (i.e., foundational, high-level language). In the future, this tool can be used to identify children who are at risk for academic failure. The study will provide preliminary data to assist in the process of validating a much-needed tool for this purpose.

6. Explain how the benefits of this research outweigh the potential risks and how these risks are justified.

Risks are minimal to subjects; potential benefit of developing an improved assessment tool justifies study.

I. HUMAN RESEARCH SUBJECTS

Subject Population:

Anticipated number of subjects participating on a unit average: 0

Anticipated number of subjects participating on an off-campus site: 15

What is the gender of the subjects? Male Female Both

What is the subject age category? Minor (less than 18 years) Adult (older than 18 years) Both

Please describe the subject age group as specifically as possible. (Example: minors, male and female high school students, 14-17 yrs. or adults, females, 50-60 yrs.)

Preschool children (both males and females) between the ages of 3-years 5-months to 5-years 6-months of age.

1. To what health category will the human subjects belong? (Example: general population, healthy adult, children with histories of depression)

Children who are developing typically

2. What will be the total duration of involvement for each subject in the study? (Example: one, 15-minute interview, three, 10-minute surveys, one questionnaire, approximate completion time: 1 hour)

Approximate completion time: 50 minutes per child. (two 10-minute assessments, one 30 minute assessment)

3. Is the research limited to any particular age, gender, ethnic, or racial group? (If an equitable recruitment from all sub-populations is not anticipated, please provide justification for weighted targeted sampling).

The current study is limited to children who are English-language speakers. At this time the tool is designed to be used with children who use English as a second (or non-primary) language. Children above 3-years 0-months will also not be included due to the study’s focus on assessment of young children. For this reason, the ages are limited to children 3-years 0-months to 5-years 6-months of age who are typically developing (i.e., do not have any developmental delays as identified by preschool teachers).

4. Will any of the following vulnerable populations be targeted for subject recruitment?

- Minor
- Minority
- Pregnant women
- Prisoners
- Mentally Incapacitated
- Terminally Ill
- Non-English speaking
- Elderly
- Severely Psychologically Disordered
- UT Students

5. What safeguards are in place to protect vulnerable populations involved within the proposed research?

Assessments will take place in the children’s preschool environment (familiar location) and the child’s teacher may be present (if the child requests). The evaluators are trained in working with young children.
8. Outline the criteria for selection and exclusion of subjects.

Children will be selected based on parent's approval. An information sheet (see attached) and permission form will be sent home with any eligible children (2 years 5 months to 5 years 6 months of age, developing typically, English as a primary language). If more than 15 permission forms are received, 15 children will be randomly selected from the total pool of children with parental permission.

7. Will subjects receive compensation for their participation, monetary or otherwise? Yes ☐ No ☒
   If yes, please specify. NONE.

8. What financial obligations will subjects incur as a result of participating in this research? Identify expenses such as travel costs, parking fees, missed work, etc. Please be as specific as possible.
   NONE.

---

J. RECRUITMENT PROCEDURES

1. What method(s) will be used to identify and recruit prospective subjects? Specify the source of potential subjects.

   Flyers and letters of consent will be sent to parents of eligible children in their home based on teacher recommendation. Children will be recruited from the All God's Children Christian Preschool.

2. Check all types of recruitment materials that will be utilized in the study.

   - ☒ Flyers/posters
   - ☐ Brochures
   - ☐ Newsletters
   - ☐ Internet
   - ☐ Radio
   - ☐ Contact letters to students or clients
   - ☐ Other (Describe): E-mail
   - ☐ None

   Attach copies of all recruitment materials to this application.

3. Will you access existing stored data, records, etc. for your recruitment purposes? If yes, specify the sources.
   NO
K. INFORMED CONSENT AND ASSENT

Select only one of the three boxes below:

☐ I am attaching a copy of all Consent and Assent forms that will be used in this study and will answer the questions below. A letter of consent is generally required from all adult research participants unless specifically waived by the IRB. A letter of assent is required of all minor research participants (age 8-17) unless specifically waived by the IRB.

The IRB has provided a template containing the elements of informed consent (per 45 CFR 116) at the UT-DHRRIRB website http://www.utdallas.edu/research/admin/irbtemplate.html. Using the template is strongly suggested in order to eliminate errors and revisions.

☐ I am requesting that the IRB waive my requirement for a signed letter of consent/assent. A signed consent form may be waived if the IRB finds either:

1. That the only record linking the subject and the research would be the consent document and the principal risk would be potential harm resulting from a breach of confidentiality. Each subject will be asked whether the subject wants documentation linking the subject with the research, and the subject’s wishes will govern;
2. That the research presents no more than minimal risk to subjects and involves no procedures for which written consent is normally required outside of the research context.

State specifically why you are asking the IRB to waive the requirement for a signed letter of consent/assent.

Answer all of the Questions Below

1. How and where will informed consent/assent be obtained? (e.g., in the school, investigator’s office, etc.)

The consent forms will be sent home in the student’s backpack to the parents of all eligible children.

2. Will there be an opportunity for potential subjects to take the consent form home to consider the options and to discuss participation with family members? If not, explain why.

Yes, see above.

3. If subjects are minors or mentally disabled, describe how and from whom permission will be granted?

Finger printing the study and consent form will be sent to the parents of all eligible participants.
5. How and by whom will it be determined that the subjects or their legal representative understand the research project and their rights as participants?

The child's teacher will indicate to the researcher if the parents are non-English speakers, or if they state the parents would not understand the written information. Those parents will not be contacted for the study.

6. Where will the record(s) of consent/assent be stored?

In the principal investigator's office in a locked filing cabinet.

---

**L. CONFIDENTIALITY OF INFORMATION COLLECTED**

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What are the methods used to ensure confidentiality of participation?</td>
<td>Students will be assigned an ID number.</td>
</tr>
<tr>
<td>2. How will data be collected and recorded?</td>
<td>Data forms will be coded with subject ID number (kept separate from subjects' names). All data files will be maintained with the confidential ID number only. After the data is uploaded to the secure database, the identifying information will be destroyed.</td>
</tr>
<tr>
<td>3. Where will data be stored during the study and how will they be secured?</td>
<td>In locked cabinet in the principal investigator's office.</td>
</tr>
<tr>
<td>4. Who will have access to the data and/or to coding schemes?</td>
<td>Principal investigator.</td>
</tr>
<tr>
<td>5. If data with identifiers will be released, specify the person(s) or agency to whom this information will be released?</td>
<td>Data will not be released.</td>
</tr>
<tr>
<td>6. What will happen to the raw data when the research is completed?</td>
<td>The raw data will be destroyed after 6 years.</td>
</tr>
</tbody>
</table>

---

**M. ASSURANCES - Principal Investigator or Faculty Advisor AND Student Investigator:**

We certify that the information provided in this application is complete and accurate.

- We understand that as the Principal Investigator or Faculty Advisor AND the Student Investigator, we have responsibility for the protection of the rights and welfare of human subjects and the strict adherence to any study-specific requirements imposed by the IRB.
- We agree to comply with all IRB and institutional policies and procedures, as well as with all applicable federal, state, and local laws and regulations regarding the protection of human subjects in research and the conduct of clinical research.
- We also agree to the following:

  1. to accept responsibility for the scientific and ethical conduct of this research study,
  2. to obtain prior approval from the Institutional Review Board before enrolling or altering the research protocol or implementing changes in the approved protocol/protocol form, study sites, or study personnel, recruitment procedures,
  3. to immediately report to the IRB any serious adverse events and/or unanticipated effects on subjects which may occur as a result of this study.
4. To train study personnel in the proper conduct of human subjects research.
6. To complete the Continuing Review and Final Report Forms required by the UT SEU IRB.

- We, the hereby attest that all named study personnel have read the protocol, understood the study, and are fully knowledgeable of all details of the protocol and are able to answer all questions from research subjects such as risks and alternatives. The above named study personnel may obtain informed consent from research subjects.

<table>
<thead>
<tr>
<th>Signature of Principal Investigator / Faculty Advisor</th>
<th>Date</th>
<th>Signature of Student Investigator</th>
<th>Date</th>
</tr>
</thead>
</table>

N. APPLICATION ENCLOSURES CHECKLIST - PLEASE REVIEW THE FOLLOWING CHECKLIST FOR APPLICATION COMPLETENESS. IRB APPLICATIONS THAT ARE NOT COMPLETE OR DO NOT HAVE THE APPROPRIATE SIGNATURES OR ATTACHMENTS WILL BE RETURNED FOR RESUBMISSION.

Check all items that are included in your submission for initial review.

The following must be included in the submission for initial review:

- [ ] IRB Application, with all of the necessary signatures (Sections III Assurance)
- [ ] Human Subjects Research Training Forms (for all study personnel who have not yet submitted the form) ALREADY ON FILE
- [ ] UT Faculty/Staff Conflict-of-Interest Form for Sponsored or Un-sponsored Human Research

Include the following (if applicable):

- [ ] Current IRB Approval letters from other sites with IRBs (Section C.)
- [ ] Approval letters from all sites where research will be conducted. (Section G.)
- [ ] Recruitment Information (Section I: Ads, Web postings, letters etc.)
- [ ] Consent/Assent forms, if applicable. (Section H.)
- [ ] Additional Information PI considers important for review by UT SEU IRB.
1. Risks to subjects are minimized, a) by the use of procedures consistent with sound research design which do not expose subjects to unnecessary risk, and b) when appropriate, by the use of procedures already being performed on the subjects for diagnostic or treatment purposes.

2. Risks to subjects are reasonable in relation to any benefits that might be expected from taking part in a research study and to the importance of the knowledge that may result.

3. Selection of subjects is fair and equitable. For example, the UT SBIR IRB seeks to determine that no eligible individuals are denied the opportunity to take part in any study, particularly those from which they may benefit, based on arbitrary criteria such as sex, age, or social or economic status.

4. Participation is voluntary and informed consent must be obtained from each prospective subject, or, when appropriate, from the subject's legally authorized representative in accordance with 45 CFR 46.116 and 21 CFR 81.116.

5. Informed consent must be appropriately documented, in accordance with, and to the extent required by 45 CFR 46.117 and 21 CFR 81.117.

6. Research interventions have been separately identified and assessed from those interventions that are therapeutic.

7. When appropriate, the research plan provides for monitoring the data collected to protect the safety of subjects.

8. When appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data.

9. When appropriate, there are additional safeguards in place to protect the rights and welfare of subjects that are likely to be members of a vulnerable population.