WH- QUESTIONS AND INDIVIDUALS WITH INTELLECTUAL DISABILITY

Eric Sanders

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Allied Health Sciences, Division of Speech and Hearing Sciences

Chapel Hill
2012

Approved by:

Karen Erickson, Ph. D.

Elizabeth Crais, Ph. D.

David Koppenhaver, Ph. D.

Patsy Pierce, Ph. D.

Linda Watson, Ph. D.

ABSTRACT

ERIC SANDERS: Wh- Questions and Individuals with Intellectual Disability
(Under the direction of Karen Erickson)

The purpose of this study was to investigate the comprehension of questions beginning with different wh- question words presented in two referential conditions to individuals with intellectual disability (ID). Thirty-nine school-age participants completed a battery of *who, what, where, when, why,* and *how* questions with and without a picture referent. Students generally answered more questions correctly in the no referent condition. Across conditions the rank ordering of correct responses for each question word was: *what, who, where, how, why, when*. While logistic regression analysis revealed a finding of no overall statistical significance between the probability of answering questions correctly between referential conditions, there were important qualitative differences in the way students answered certain wh- questions based on referential condition.

The questions presented were grouped and compared based on conceptual complexity with *who, what, where* categorized as concrete and *when, why, how* as abstract. Participants answered 89% of concrete and 56% of abstract questions correctly across both conditions. A statistically significant difference between the probability of answering correctly concrete and abstract questions was found. There was also a statistically significant relationship between receptive vocabulary and the probability of answering all questions correctly.

Additionally, logistic regression models revealed that receptive vocabulary ability may be more related to answering abstract questions than concrete question.

Analysis of incorrect answers showed that 85% of errors could be meaningfully categorized. Of these responses, 25% were categorically related to the particular wh-question word that was targeted. Further, 10% would have been correct if a different question word was substituted for the target. Additionally, 27% of the responses were related to the topic of the question and the remaining were either "I don't know/No responses", repetitions, or unrelated. The categories were generally the same across conditions. Students did, however, produce more of these incorrect responses that were related to the topic in the picture condition. The results of this study point to the importance of learning more about ways to support question comprehension and answering to help individuals with ID answer all types of questions more effectively.

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support of many people. Advisors, friends, family, and many others helped provide the support necessary to help see me through this entire process. I will forever be grateful for having these people in my life during this time.

The first group of people I would like to thank are the people who were responsible for and oversaw my doctoral training. Drs. Betsy Crais, David Koppenhaver, Patsy Pierce, and Linda Watson all had vital roles in this dissertation. These members of my dissertation committee put in many hours of work over the last few years consulting with me about the project, helping shape my ideas, and providing important feedback. I have relied on each of these professors for advice not only about my dissertation, but also about important professional and life issues. Without their wisdom and counsel, this dissertation would never have been completed.

I would also like to thank my mentor, Karen Erickson, for her work guiding me through this process for the past five years. She was the reason I decided to pursue my doctorate at the University of North Carolina and the Center for Literacy and Disability Studies. Her devotion to individuals with disabilities is unparalleled and inspires me to continue to do work that makes a difference their lives. I am so appreciative of the effort she has made working with me through this entire process. Through working with her I have learned valuable lessons about dedication and hard work that I will take with me for the rest of my career.

I have been fortunate to work at the Center for Literacy and Disability Studies for the past five years. The mission of the Center and the hard work of its faculty continually inspire my work to help individuals with disabilities. I would like to thank Dr. David Yoder and Gretchen Hanser for their commitment to bettering the lives of people with the most significant disabilities and for always listening and sharing advice. I am also appreciative for my friendship with Kristin Nellenbach and Skip Ryan and all of the support they have given me. Finally, I would like to thank Penny Hatch for helping me to navigate this process. The support of the Center has been instrumental in my ability to make it through this course of study.

Another group of people that have been important to me throughout this process are my fellow students. Many of these students have become friends that I will have throughout my life. I am greatly appreciative to Kaitlyn Wilson, Jessica Dykstra, Jessie Klusek, Angela Bonino, Hillary Harper, and Dwight Irwin for their support. There is no way this process could have been completed without this support system. I would especially like to thank Thomas Page for his support and friendship over the last thirteen years. He has helped me in innumerable ways from helping me find a place to live to working with me through different ideas. Thank you.

I have also been lucky enough to have met Emily Dings during my time here at UNC. She has been unbelievably supportive throughout this process. I genuinely appreciate all she has done for me, often at the sacrifice of her own time. There is absolutely no way I could have completed this dissertation without her support. I am truly thankful for everything she has given me and am grateful that she is in my life.

Last, I would like to thank my family. My parents instilled a work ethic within me as a child that has made the completion of this dissertation possible. Additionally, I learned about kindness and the importance of doing the right thing. This has colored and will continue to color my work. I would also like to thank my sister, Robin, brother-in-law, Terry and my nephew, Eli. My entire family has believed in me throughout this process. This belief has allowed me to continue pushing through, even when things were difficult. I am thankful for all of their support throughout the years.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

Statement of the Problem

The ability to answer questions has ramifications on language, literacy, and academic development for individuals from birth through adulthood. During childhood, early experiences with informational exchanges typically happen through question-answer exchanges (Parnell & Amerman, 1983). Early experiences answering different types of questions are important in that they provide occasions for young children to develop language through the production of different types of responses (de Rivera, Girolametto & Weitzman, 2005). In preschool, the questions teachers ask are important because they teach children about decontextualized language that is vital to be able to read with comprehension and understand academic discourse (Massey, Pence, Justice, & Bowles, 2008). Throughout the school years, answering questions is integral to the continued development of literacy skills and achievement of academic success (e.g., Kintsch, 2005).

The development of question-answering skills has been examined from a number of perspectives in individuals who are typically developing. For example, a developmental sequence of wh-question word acquisition in receptive (e.g., Tyack & Ingram, 1977) and expressive (e.g., Bloom, Merkin, & Wooten, 1982) language has been identified. Factors that improve comprehension of different types of questions (e.g., referential conditions) have been identified (e.g., Parnell, Patterson, Harding, 1984). However, there is no research that has investigated the same range of questions and factors related to comprehension in individuals with intellectual disabilities (ID).

Given its clear importance in language and academic development, it is surprising that question comprehension has not been more thoroughly investigated in students with ID. The ability to answer different types of questions involves the integration of conceptual knowledge (Owens, 2008) and linguistic skills (e.g., Roeper, 2004). The fact that individuals with ID have likely deficits in one or more domains of language (e.g., Rosenberg & Abbeduto, 1993; Paul, 2007) suggests that they would have difficulty understanding and answering questions, but the issue has not been systematically investigated.

A logical place to initiate a line of research regarding question answering for individuals with ID is to explore the relative difficulty of different question types. Further, research on question comprehension has revealed that there are differences in the way individuals who are typically developing and those with language impairment answer questions based on the way they are presented (Parnell, Amerman, & Harting, 1986; Parnell et al., 1984). Understanding if and how similar differences impact question comprehension in individuals with ID has important intervention implications.

Determining ways to improve question comprehension is paramount to giving students with ID an opportunity to develop language, literacy, and academic skills. This requires learning more about the questions individuals with ID understand. It also requires understanding more about the conditions that may help them answer questions more successfully. Improving these understandings is a first step in developing evidence-based interventions that may aid in the development of question comprehension in the future.

Purpose

The purpose of this study was to investigate the question-answering abilities of students with ID across different conditions. This study examined the comprehension of

questions that begin with the following wh- question words: *who*, *what*, *where*, *when*, *why*, and *how*. These question words have been commonly targeted in developmental studies of question comprehension (e.g., Ervin-Tripp, 1970; Tyack & Ingram, 1977) and are commonly used in academic settings (see Common Core State Standards, 2010). Additionally, questions were asked in two distinct referential conditions in an attempt to determine if a picture referent could support individuals with ID in understanding and answering questions. Questions were also divided into groups based on their relative concreteness and abstractness and compared to determine if there were differences in the way they were comprehended. Finally, this study was designed to investigate the relationship between receptive vocabulary and question answering in individuals with ID. The information gained from this research helps to paint a more complete picture of question-answering ability in this population.

Varying Levels of Wh- Questions

Wh- questions vary in terms of their syntactic, semantic, and conceptual demands. The results of a handful of studies suggest that individuals with ID respond differently to questions based on their conceptual level (e.g., Zetlin & Gallimore, 1983; Hewitt, 1998), but these differences have never been invested with reference to the specific wh- words used. Yet, there is evidence that wh-question words vary in their conceptual complexity (e.g., Clancy, 1989). In this study, the comprehension of questions with different wh- question words was systematically examined.

These differences in the conceptual complexity of different wh-question words are evident in typical development. For example, children usually comprehend *what*, *where*, and *who* before *when*, *why*, and *how*. Although there are sometimes variations of this order, early in development it is easy to conceptualize these types of questions as being grouped together.

For example, *who*, *what*, and *where* often function like pronouns in that they replace a specific sentence constituent. Further, these usually pertain to something concrete like an object or a thing. In fact, *what* and *where* are often used by parents when speaking with infants because they are about something that can be pointed to and are therefore good for early language tutoring and informational exchange (Owens, 2008). *When*, *why*, and *how* usually pertain to the semantic relations of the words in the sentence they refer to and ask for additional information (*where* can also function this way) (Bloom et al., 1982). These questions words are also different from *what*, *where*, and *who* conceptually. That is, in order to comprehend these questions children need to have some concept of time, causality, and means/manner respectively.

Although developmental studies have specifically looked at these aspects of question answering, the only differentiation between questions in studies with individuals with ID has been in terms of general conceptual level (e.g., low versus high levels of questions). It is possible that school-age individuals with ID do not yet understand specific question words and this confounds the difficulty they have with more conceptually complex question types. By investigating the difference between these basic types of question words specifically, it may be possible to understand more about what students know about the questions they are asked. This would in turn allow for greater specificity of question targets in terms of educational and therapeutic intervention planning.

The Influence of Referential Condition on Question Answering

The ability to answer questions may be impacted by the availability of the referential source of that question (Parnell & Amerman, 1983). Age-related differences have been found in the way typically developing children understand questions based on whether a referent

for the question was present or not (Parnell et al., 1984). Children with language impairments appear to be even more sensitive than typically developing children to the presence of a referent (Parnell et al., 1986).

Referential source is important to consider in studies of question answering in individuals with ID because it is reasonable to expect that individuals with ID will benefit from the use of picture referents. This assumption is based on the fact that children with ID may experience difficulties regarding question referents in a manner that is similar to children with language impairment. Furthermore, pictures have traditionally been used successfully in interventions to teach children with ID how to answer certain types of questions (e.g., Jahr, 2001; Hundert & Delft, 2009). By learning more about whether individuals with ID understand questions better in different referential conditions, it may be possible that better interventions can ultimately be designed to help these students answer questions more effectively.

Incorrect Answers

Investigating what types of questions individuals with ID can answer is important, but so is understanding the nature of their incorrect responses. Valuable insight into what an individual understands about a question can be obtained through analyzing incorrect answers. These errors have been analyzed in attempts to determine what strategies children use when they do not understand a question (e.g., Ervin-Tripp, 1970, Tyack & Ingram, 1977). Additionally, error analysis has been used to explore whether children know something about the question form even though their answer is incorrect (Cairns & Hsu, 1978; Parnell et al., 1984). In these studies, coding schemes have been utilized to determine whether a response has been categorically appropriate, but incorrect (e.g., giving an incorrect time response for a

when question). This type of analysis has revealed that children's incorrect responses vary in a particular developmental manner as children move toward correct responses. For example, children have been shown to answer with categorically appropriate answers before they answer questions correctly (Parnell et al., 1984). Further, this same type of error analysis has revealed that individuals with language impairment may also progress through similar developmental stages as they learn to answer questions correctly (Lee & Ashmore, 1984; Parnell et al., 1986). Identifying whether there are similar trends in individuals with ID may help interventionists and/or teachers who work with individuals with ID.

Receptive Vocabulary and Question Understanding

One of the difficulties in studying question comprehension in ID is the fact that language skills can be highly variable. Individuals with ID typically experience difficulties across domains of language (e.g., McDuffie & Abbeduto, 2009). Distinct profiles of linguistic ability in ID are known to be highly variable, even when the genetic etiology of the impairment is known (Paul, 2007). Because of this variability, it is important to determine if language skills are related to question answering.

One specific domain of language that should relate to the impact of specific wh-question words on question comprehension is semantics. Wh-question words are semantic units with distinct meanings. Therefore, it is possible that understanding a variety of question words may relate to overall understanding of receptive vocabulary. Because of the linguistic heterogeneity in individuals with ID, beginning with receptive vocabulary is a logical place to start investigating how differences in domains of language may impact the comprehension and answering of wh- questions.

Summary

Question answering is important to linguistic and educational development. The current study was designed to fill gaps in the current literature. First and foremost, it addressed the fact that no research currently exists regarding the impact of wh-question words and picture referents on question comprehension for individuals with ID. Building on the work with individuals with and without language impairments (e.g., Lee & Ashmore, 1983; Tyack & Ingram, 1977), the investigation probed the following question words: *who*, *what*, *where*, *when*, *why*, and *how*. Furthermore, it explored the relative ease of answering these questions based on conceptual complexity (e.g., Tyack & Ingram, 1977). Based on the fact that picture referents appear to support question comprehension for at least some individuals (Parnell et al.,1984), the study compared question comprehension in a picture and no-picture condition. Given the role that language, in particular semantics, plays in question comprehension, the relationship between receptive vocabulary and question comprehension was explored. Finally, given our understanding that the errors individuals make when answering questions are often meaningful (Parnell & Amerman, 1983), incorrect responses were analyzed to determine if any patterns emerged.

This work is necessary to learn more about the types of wh-question words individuals with ID understand. This study is a first step in identifying the wh- questions that individuals with ID may have strengths or weaknesses answering. The information gained from this study will help guide educators and clinicians who work with these individuals to better understand questions.

CHAPTER 2

Review of the Literature

The ability to answer questions provides a window through which to view what an individual knows or does not know about a given topic. Answering questions is integrally related to the development of language (e.g., Bloom et al., 1982), literacy (e.g., National Reading Panel, 2000), and academic success (e.g., Dillon, 1988). In typical development, there is much known about the way questioning understanding and answering emerge; however, little is known about this development in individuals who have intellectual disabilities (ID). An investigation of these skills in individuals with ID requires an understanding of the extant literature regarding the relationship between language and the development of question comprehension, the development of question comprehension across individuals with and without disabilities, and the language development of individuals with ID.

The Domains of Language and Their Relation to Wh- Question Comprehension

Comprehending questions is a complex task that involves the coordination of a number of linguistic and cognitive skills. It is difficult to parse out the language skills involved in question comprehension because most of these skills interact and affect one another at some level. Nonetheless, the current investigation considered the domains of semantics and syntax in the design of the measures and analysis of the results. As such, each of these domains of language is described below as they relate to the comprehension of questions.

Semantics and Their Relation to Question Comprehension

Semantic abilities play a pivotal role in the ability to understand questions. Most prominently, individuals must understand the meaning of wh- question words (e.g., *who*, *what*, *where*, *when*, *why*, *how*). For all wh- question words except *how*, the phonetic cue (/wh/) at the beginning of the word operates as a signal that indicates that the word belongs to a distinct grammatical category (DeVilliers, Roeper, Bland-Stewart, & Pearson, 2009; Radford, 2009). The phonetic cue, along with its position in the beginning of the sentence, typically signals listeners that they are being asked a question.

Wh- words generally stand for a missing constituent (i.e., a syntactic structure) that the respondednt is asked to provide. The information sought when someone asks a wh-question varies based on the wh- word. *What* typically stands for a thing, *who* for a person, and *where* for a place or direction (Winzemer, 1980; De Villiers et al., 2009). These question words often function like pronouns for the constituent they replace (Owens 2008). *When* stands for a particular time and *why* for a causal reason. *How* questions can inquire about the means (e.g., "*How* did you eat your spaghetti?"—"With a fork.") or the manner in which something is done ("*How* did he ride?" — "Quickly.") (Winzemer, 1980; De Villiers et al., 2009).

Wh- question words vary in terms of the abstractness of the constituent they stand for. For example, *who* or *what* questions typically refer to tangible nouns that can often be seen in the immediate environment. Answering *why* and *when* questions, however, requires more complex reasoning involving concepts like causality and time. For this reason, it has been suggested that individuals need to understand the presupposed concept a wh- word stands for before it can be correctly understood (Clancy, 1989).

Theories about the link between conceptual knowledge and the understanding of question words are equivocal because of a lack of evidence correlating the mastery of concepts to an understanding of particular question words (Clancy, 1989). Related to this, children use other linguistic structures related to the same conceptual basis as words like *when* and *why* before asking questions that begin with those words. For example, because is often used appropriately before children begin asking *why* questions. Also, before *when* questions are asked, connectives that require temporal concepts such as and and then are typically used (Bloom et al., 1982).

It should be noted that wh- question words can have different functions that are more conceptually complex. Teachers often use these types of questions in academic discourse (Hunkins, 1989). For example, *what* questions have functions so they can serve as causal antecedents (e.g., "*What* caused the Dolphins to lose the game?"), causal consequences (e.g., "*What* happened to the restaurant after the fire?"), or judgments (e.g., "*What* do you think of the way the president handled the economy?"). Likewise, *why* has additional functions. The expression of cause as an answer to *why* questions can be further divided into whether the answerer is asked to explain the reasons or goals of an action (e.g., "*Why* did the boy decide to ride his bicycle home?") or to reflect personal expectations based on given information (e.g., "*Why* would you argue for the debt ceiling to be raised?"). Even *when* questions can ask students about the understanding of conventions as opposed to strictly being concerned with time identification (e.g., "*When* are you supposed to raise your hand?"). Outside of manner and means, *how* can also signal a quantification response (e.g., "*How* many dollars do you have?") (Lehnert, 1978). In the current investigation, the functions of the question words included were defined and matched across context.

Developmentally, the use of different question words appears to be influenced by the relationship between the question words themselves and the complexity of language used by caregivers (Bloom et al., 1982; Rowland, Pine, Lieven, & Theakston, 2003). For example, it has been shown that question words like *who* and *what* are used more by parents when speaking to younger children because a referent may be immediately present. It is speculated that children use those words first because of that input and because the verbs or auxiliary verbs that are used with those question words are simpler, and therefore easier for children to understand than those that may be used with question words like *when* and *how* (Bloom et al., 1982). Additionally, it has been speculated that parents may only use certain question words when they feel as though their children can conceptually understand them (Clancy, 1989; Rowland et al., 2003).

Comprehension of wh- questions involves more than just understanding the wh-word itself. Obviously, unless words can be determined from context, individuals need to understand each of the words in a question to comprehend the overall meaning of the question. Even if the individual words are understood, the semantic features of words such as the verb in a wh- question may influence a response. For example, verbs like "help" have been shown to elicit causal responses more frequently than the verb "touch" when paired with the question word *why*. It has been speculated that this might be because "touch" primes children to think of an object or location rather than a cause (Tyack & Ingram, 1977; Winzemer, 1980). These semantic characteristics make it difficult to parse out true understanding of wh- words and should be considered when interpreting the results of studies designed to assess comprehension of wh- question words.

Syntax and Its Relation to Question Comprehension

Semantic skills are necessary to be able to answer questions appropriately, but knowledge of the words is insufficient because questions can vary in meaning even if they have the same words. For example, the questions "*Who* is hitting the boy?" and "*Who* is the boy hitting?" ask for different information although they are composed of the same words. Like other linguistic structures, syntactic rules are used to bind words and morphemes together to create the form that is used to comprehend and express a message. Wh- questions are noteworthy for their syntactic complexity and abstractness (e.g., de Villiers et al., 2009).

The syntactic structure of wh- questions ranges from those that have a single clause (e.g., "*What* is the dog eating?"), to those with embedded wh- words (e.g., "Do you know *what*'s for lunch?), multiple wh- words (e.g., "*Who* knows *when* the game is?"), and multiple clauses (e.g., "*What* will the dog eat after getting back from the vet?"). One-clause questions are relatively "simple" whereas multiple-clause questions are more complex.

The relationship between the wh- question words themselves and syntax plays an important role in an individual's ability to comprehend questions. Although question words replace constituents, there is syntactic complexity related to what constituents they replace. It is useful to think of questions as being either argument or adjunct questions. In argument questions, the question word refers to the argument of the verb. These include subjects and objects of a sentence or possibly indirect objects of verbs. These are obligatory in that a sentence is incomplete without them. *Who* and *what* are argument questions because they refer to information that must be in the sentence. *Where* is occasionally obligatory because verbs like "put" require an argument. Adjunct questions are asking for additional information. These typically include the question words *when*, *why*, and *how*. Take the

sentence, “Tom put the ball in the closet.” In this case, the question words *who*, *what*, and *where* all refer to something in the sentence that has to be there for the sentence to be complete. To answer *when*, *why*, and *how* questions about that sentence, the answerer must supply adjunct or additional information (de Villiers et al., 2009).

Recently, linguists have shown particular interest in the production and comprehension of two different types of simple wh- questions: those referred to as subject and object questions. In subject questions, the wh- word represents the subject of the sentence. For example, the answer to the question “*Who* touched the boy?” could be the “the girl.” Here, *who* takes the place of the constituent “the girl” in the sentence “The girl touched the boy.” The typical subject-verb-object order seen in English phrases is not disturbed in this case. Conversely, in object questions the constituent replaces the object position of the sentence. For example, the answer to the question “*Who* did the boy touch?” could be “the girl.” In this case, “the girl” is in the object position of the sentence “The boy touched the girl.” Additionally, wh- object questions require the inversion of the subject and auxiliary in the question whereas subject questions do not. This deviates from the subject-verb-object order typically seen in English phrases. *Who*, *what*, and which wh- words can be used to ask both subject and object questions.

When answering object questions a gap exists between the noun phrase and the object (the answer), which does not exist in a subject question. It is speculated that this gap may make these questions more difficult for children to ask and answer (O’Grady, 1997). Indeed, it has been found that infants can understand simple subject questions at fifteen months, but do not understand simple object questions until twenty months (Seidl, Hollich, & Juczyk, 2003). However, research on the comprehension of these types of questions in older children

has shown equivocal results (e.g., Stromswold, 1995) and suggests that differences may be influenced by the question word used as opposed to the syntax of the entire question (Tyack & Ingram, 1977). In the current investigation, all of the questions were object questions to insure that differences were the result of the relative complexity of the wh- question word rather than the difference between subject and object questions.

Wh- questions vary in a number of ways. Syntactic complexity is just one of those. Syntactic complexity may influence whether a question is comprehended above and beyond the influence of the wh-question word that is used. For this reason, to determine whether individuals understand question words, the syntactic complexity of the question itself needs to be controlled, or, in the very least, considered.

Developmental Studies of Wh- Question Word Comprehension

Many attempts have been made to determine the developmental sequence of wh-question word production in typical development. There are, however, significantly fewer studies of the development of question word comprehension. This deficit is surprising in that it is generally acknowledged that children do not use question forms appropriately (e.g., not in a formulaic way) until they comprehend them (Bloom & Lahey, 1978; Brown, 1968; Parnell & Amerman, 1983). The development of comprehension of questions with wh-words has primarily been measured in two ways. The first is through longitudinal, fine-grained linguistic analysis of interactions between mother-child dyads in a variety of contexts (e.g., storybook reading or playing with toys). Although these have provided valuable information, it is difficult to generalize results because there are typically only a few participants (e.g., Brown, 1968; Clancy, 1989). Further, the types of questions that are asked vary in terms of length, vocabulary, and context. The second method of investigation has

been experimental, cross-sectional studies that compare groups of children at different ages to determine their relative success in answering a variety of wh- questions (e.g., Ervin-Tripp, 1970, Tyack & Ingram, 1977).

Studies of Wh- Question Word Comprehension

Roger Brown (1968) conducted one of the first studies investigating the development of wh- question word comprehension. He compared the production and comprehension of questions in three preschool-aged children. He found that all three children were able to answer questions with a variety of wh- words before they were able to competently produce them. The one exception to this rule was in the comprehension of *why* questions. One child produced *why* questions before he was able to answer them appropriately. The other children followed the typical pattern of consistently producing questions only after they were able to respond appropriately. None of the children responded to *why* questions appropriately before they had a mean length of utterance (MLU) of 2.75 words.

Brown did not report specific information regarding the ages when the children displayed their understanding of the different wh- question words, but he did note that *when*, *how*, and *why* questions were produced after *who*, *what*, and *where* questions. Furthermore, by the time the children had an MLU of 2.75 words, they were correctly producing a variety of sentence constituents (e.g., noun-phrase subjects, main verbs, noun-phrase objects, locative adverbials) in declaratives that they were not producing before this point. Subsequently, they were answering questions that required these constituents correctly about half of the time.

Ervin-Tripp (1970) was the first researcher to design a study specifically intended to determine the developmental comprehension of wh- words through experimental means. She

observed five children and recorded their answers to certain types of questions and then interviewed twenty-four children (ages 2;6 to 4;2 years) once a month and asked standard questions about a children's book. Ervin-Tripp (1970) references the extreme variability between the participants, but she also pieced together an order of acquisition for the way they answered questions over this time period. To determine whether the participants comprehended the question, responses were judged in terms of their category agreement. A reply to a question was determined to have category agreement if it matched the adult expectation of the appropriate grammatical category and semantic feature (e.g., the answer to a *where* question should be an adverbial locative). The order that emerged from the combination of both studies is as follows: *what, where, what*-do, *whose, who, why, where*-from, *how,* and *when*. There was a difference between the comprehension of early *who*-subject and *who*-object questions, and some difficulties were observed in older children with *who*-object questions. In describing the results, Ervin-Tripp (1970) highlighted the need for more syntactically controlled questions as stimuli in future studies.

In the next wide-scale study examining wh- question word comprehension in children who are typically developing, Tyack and Ingram (1977) designed a paradigm not only intended to look at wh- question word comprehension, but also at the influence of the verb and whether there was a difference between the comprehension of simple subject and object questions. They assessed 100 children who ranged in age from 3;0 to 5;5 and placed them into five age groups that were divided into six-month intervals. Their targets were the comprehension of *who*-object, *who*-subject, *what*-object, *what*, subject, *where*, *how*, *why*, and *when* questions (all object). Unlike Ervin-Tripp (1970), they used syntactic frames to control for the syntax of the questions. They controlled the semantics of the questions by only using

verbs from the first 1,000 words in the *Thorndike-Lorge Book of 30,000 Words* (Thorndike & Lorge, 1944) and included an equal number of frames with transitive and intransitive verbs. To elicit responses from the children, they asked questions paired with staged photographs.

Similar to Ervin-Tripp's (1970) category-agreement coding scheme, broad categories were used to determine the accuracy of responses (e.g. a person for *who*). Overall, the number of correct responses to questions increased as a function of age. The percentage of correct responses also varied on the basis of the syntactic frame, type of verb, and question word. When the data was aggregated across all age groups the following order of correct answers per question word, verb type, and syntactic frame were as follows: (1) *Where* – intransitive, (2) *why* – intransitive, (3) *why* – transitive, (4) *who* – subject, (5) *where* – transitive, (6) *what* – object,(7) *who* – object, (8) *when* – intransitive, (9) *when* – transitive, (10) *how* – transitive, (11) *how* – intransitive, (12) *what* – subject. As a whole, it is difficult to parse out a developmental sequence of wh- question words in this study because of the role of the verbs and syntactic frame of the question. For example, it is difficult to ascertain whether children understood simple wh- subject questions before wh- object questions. Tyack and Ingram (1977) used syntactic frames that to do this. These did not have the appearance of being natural (e.g., "*What* is riding the boy?"). The current investigation was designed to address this challenge by investigating only simple wh- object questions while keeping them as natural as possible.

Cairns and Hsu (1978) also conducted a study that explored the order of acquisition of wh- question words in fifty children between the ages of 3;0 and 5;6 divided into groups five subgroups in six-month intervals. They showed the children videotaped vignettes of a family and asked *who*, *when*, *why*, and *how* questions after each. They included three types

of *who* questions: *who*-object with do support (e.g., "*Who* did the boy feed?), *who*-object with progressive aspect (e.g., "*Who* was daddy feeding?"), *who*-subject (e.g., "*Who* hugged the boy?"), as well as *why*, *when*, and *how* questions. Significant differences were found between all age groups on a combination of all questions (with the exception of the 4;0-4;5 group versus the 4;5-5;0 group). The following order of difficulty was noted: *who*-(object [do], subject, then object), *why*, *when*, *how*.

Parnell, Patterson, and Harding (1984) also conducted a study looking at the development of answers to wh-questions. Unlike in previous investigations, they studied nine different types of question forms: *what*-be, *what*-do, *where*-be, *which*-be, *who*-be, *whose*-be, *why*, *when*, and *what*-happened. They also increased the age range (3;0 to 6;11) and were the first researchers to systematically examine the influence of different referential sources on the answers that the participants gave. All other studies using experimental stimuli used a single referential source (e.g., videotaped vignettes, pictures). Here, three different referential sources were used; each one becoming progressively more abstract than the one before it. Questions were first asked about objects (Condition 1), then pictures (Condition 2), and then about something that was not present (Condition 3). The researchers did not use syntactic frames but attempted to make questions that mirrored the types of questions participants would hear at home and school. Although the majority of the questions were simple, there was great disparity in terms of their word and morpheme length. These added processing constraints may have made a difference in the comprehension of the questions.

Participants in this study were able to answer a greater proportion of questions correctly as their ages increased. In terms of referential condition, items that had no

immediate referent were the most difficult to answer except in the case of *when* questions in the object condition. Similar to the studies by Cairns and Hsu (1978) and Ervin-Tripp (1970), *what* and *where* questions were generally answered correctly more often than *why* and *when* questions. *What*-happened, however, was the most difficult question form for all of the participants to answer. As the response to this type of question is likely to be more complex than a typical *what* question, the difficulty children had answering these correctly is unsurprising.

Analyzing Answers to Questions

Almost all developmental studies of question comprehension have utilized the participant's verbal response to determine "understanding" of the question. However, the scoring criteria across these studies have not been consistent, and all studies have analyzed the incorrect responses in different ways to determine what students do when they do not provide a correct response. This results in variations that certainly influence findings regarding order of acquisition of wh- questions in typical development. Understanding the scoring criteria that have been used in extant research is important to understand the decisions made in scoring responses in the current investigation.

The original structured experimental studies of Ervin-Tripp (1970) and Tyack and Ingram (1977) that investigated wh- question comprehension analyzed answers in terms of their "category agreement." This procedure also helped the researchers describe the nature of the incorrect responses participants provided. For example, in Ervin-Tripp's (1970) work, participants generally replied with an answer that was acceptable for an already acquired wh-question form if they did not understand the question word (e.g., if the participant did not understand *why* they may have provided a *where* response). Whether they answered the

questions correctly or incorrectly, participants used a number of strategies to respond to questions including responding with (1) an appropriate response when the question word was recognized; (2) an object of the verb when there was a transitive verb in the question; (3) a causal explanation when there was an animate subject and an intransitive verb (a verb not requiring a direct object); and (4) a location or direction if it was missing when any additional intransitive verbs were included in the question.

Tyack and Ingram (1977) also categorized incorrect responses by the type of response given (e.g., if a *how* question yielded a *why* response, it was classified as a *why* interpretation). They found that the semantic features of the verb and the conceptual depth of the wh- word influence the types of incorrect responses students make. In terms of the semantic features of the verb, both "touch" and "help" are transitive verbs, but questions with the verb "touch" led to nominal or locative responses more than a causal response while questions with the verb "help" led to more causal responses regardless of the question word it was paired with.

Unlike previous studies of wh- question word comprehension that were concerned only with category agreement, Cairns and Hsu (1978) coded responses to *why*, *when*, and *how* responses based on the quality of the response. The intent was to develop a scoring scheme that would more adequately represent what the child understood about the question form. Responses were categorized into four types: Type 1 – A response indicating a failure to respond (e.g., "I don't know"); Type 2 – A response indicating some knowledge of the question type, but no more than a minimal response (e.g., "because" used to answer a *why* question); Type 3 – A response indicating an understanding of the question, but not a fully

complete answer (e.g., a semantic or syntactic error in the response); and Type 4 – A fully completely correct answer (e.g., it is adult-like with no errors).

The scoring criteria of Cairns and Hsu (1978) impacted the reported developmental order of the comprehension of different wh- words when compared to previous research. For example, *why* questions were answered less successfully than *who*. Cairns and Hsu (1978) point out that many of the answers that Tyack and Ingram (1977) would have accepted as being correct were counted as "Type 3" responses in their own scoring system. When examining the results of their scoring procedure, they found that the "rate of correctness" moved from the *why* questions (most correct) to *when*, and finally to *how* (least correct).

Parnell, Patterson, and Harding (1984) continued the trend of a conducting a more in-depth analysis of the answers that were provided. They also utilized a coding scheme designed to investigate the quality of answers rather than strictly looking at category agreement. Their scheme was based on the premise that adult listeners are most interested in the truth, logic, factualness, and credibility of children's answers. Responses were divided into two categories: one indicating "functional appropriateness" and the other "functional accuracy." Responses were judged as having functional appropriateness if they met the requirements for providing the category information required by a particular wh- word (e.g., category agreement). Functional accuracy referred to whether the actual content of the answer was right or not. Scoring was dichotomous for both categories, resulting in it either being correct or incorrect.

Some of the general trends of their results remained the same. For example, correct answers increased in terms of their functional appropriateness and accuracy as a function of age. When the scoring categories were examined separately, participants produced answers

that were more functionally appropriate than accurate. Related to age, younger children were found to give more answers that were functionally appropriate than those that were accurate when compared to older answerers. In terms of referential condition (condition 1 = object, condition 2 = picture, condition 3 = no immediate referent) and the scoring categories, items that had no immediate referent were most difficult to answer for both categories. The youngest participants (age 3) produced more appropriate than accurate responses in all referential conditions. The 4-year-old participants showed significant differences between accuracy and appropriateness in conditions 1 and 2. Participants in the 5- and 6-year-old groups did not show any difference between appropriateness and accuracy for any of the referential conditions. *When* questions were still most difficult across scoring categories and referential conditions. As in other studies, children who did not understand a particular question word gave an answer that was functionally appropriate for a question word they had already acquired.

Summary of Developmental Wh- Question Comprehension Research

The developmental studies reported here are seminal attempts at determining a developmental order of the comprehension of certain wh- question words. The exact developmental order of question-word comprehension subtly fluctuates in the studies because of methodological issues inherent in trying to measure this skill. For example, these researchers elicited responses to questions in different ways. Additionally, different scoring criteria and answer analyses were used to determine comprehension of questions. In terms of the questions themselves, the vocabulary and syntactic structure of the questions has varied across studies. Furthermore, researchers have targeted a variety of different wh- words. It should also be noted that the majority of experimental studies designed to examie wh-

question-word comprehension use a protocol that deviates from the contextual norms of a typical question-answer dialogue (Parnell & Amerman, 1983). These are concerns that should be considered in developing a protocol that measures this construct.

Although these methodological issues have been addressed in different ways with each study, it is generally understood that the developmental order of wh- question comprehension is similar to what is seen in production and that this order is relatively sound (Tager-Flusberg, 2005). With this knowledge and careful interpretation of the results of the studies, a general order of wh- question comprehension development can be inferred. Unsurprisingly, this order varies along conceptual lines. Here, children between the ages of three and four have been reported to answer questions that use *what*, *where*, and *who* given that appropriate context is provided. Later-developing question words that are answered in typical development include *when*, *how*, and *why* (Owens, 2008; Shulman & Capone, 2009; Winzemer, 1980).

Language and Intellectual Disability

Question comprehension and answering is partially dependent on language ability. Therefore, the language ability of individuals with ID is likely to have some impact on their ability to understand and answer questions. Intellectual disability is defined by limitations in both intellectual functioning and adaptive behavior that originate before the age of 18 (U.S. Department of Education, 2002). There are a wide variety of etiologies of ID, including genetic/chromosomal syndrome (e.g., Down syndrome, fragile X syndrome), asphyxiation during childbirth, and prenatal exposure to toxins (Hodapp, Griffin, Burke, Fisher, 2011; Shevell, 2008). Approximately 50% of cases of ID result from unknown causes (Shevell, 2008), although recent advances in genetic testing are making it easier to identify possible

causes of ID (Hodapp et al., 2011). ID is also known to co-occur with developmental disabilities such as autism spectrum disorders (ASD) (Matson & Shoemaker, 2009); purportedly affecting approximately 50%–70% of individuals identified as having ASD (Fombonne, 2003; Matson & Shoemaker, 2009). Regardless of the etiology of the ID, children with ID typically experience difficulties across language domains (Fowler, 1998; McDuffie & Abbeduto, 2009).

The relationship between cognition and language is not always clear-cut, as some individuals with ID show strengths and weaknesses in different domains of language relative to their cognitive abilities (Fowler, 1998). In individuals with known, genetic causes of ID, strengths and weaknesses generally appear to be linked to distinct linguistic profiles related to the phenotypic expression of that particular syndrome (Fidler, Philofsky, & Hepburn, 2007). Furthermore, there appear to be age-related differences in linguistic strengths and weaknesses within some individuals with ID (Fidler et al., 2007; McDuffie & Abbeduto, 2009). Nonetheless, language characteristics in ID are not universal and can be highly variable, even when the etiology of the impairment is known (Paul, 2007). Often, language acquisition in individuals with ID is commensurate with cognitive functioning, resulting in a delayed version of the same developmental sequence seen in typical development (Rosenberg & Abbeduto, 1993). However, this is variable (Paul, 2007), and any deficits in one or more of the domains of language (e.g., semantics, syntax, pragmatics) secondary to ID will likely lead to difficulties understanding and subsequently answering questions.

Semantic Skills and Intellectual Disability

Individuals who have difficulty with semantics have trouble using and understanding the meanings of words at the individual word, sentence, and extended-discourse level

(Fowler, 1998; McGreggor, 2009). Although individuals with ID often present with language impairment, their semantic skills are not necessarily commensurate with their level of cognitive functioning and other language skills (e.g., Roberts, Price, & Malkin, 2007). This relationship has been shown to vary as a function of the etiology, the aspect of semantics that is being investigated, and the age of the individuals studied (e.g., Chapman, 2006, Rosenberg & Abbeduto, 1993), which makes it difficult to characterize the semantic skills of individuals with ID as a whole (MacGreggor, 2009). There are, however, features related to the receptive vocabulary skills of individuals with ID that may impact the understanding of questions.

Receptive vocabulary is generally thought to be a relative strength in many individuals with ID when compared to other domains of language and mental age (Chapman, 2006; Chapman, Seung, Schwartz, et al., 1998; Facon, Facon-Bollengier, & Grubier, 2002). However, this relationship appears to be closely tied to differences in age and the type of vocabulary that is measured. For example, receptive understanding of concrete vocabulary (e.g., event and experience-based vocabulary) has been shown to be greater than receptive syntax in adolescents with Down syndrome (Chapman, 2006; Laws & Bishop, 2003) and in children with ID of mixed etiology (Facon et al., 2002). Differences in receptive vocabulary and syntax have disappeared, however, when tests examining more conceptual vocabulary (e.g., most, few, equal, high) have been employed to assess adolescents of both of these populations (Chapman, 2006). Therefore, it has been hypothesized that strengths in more concrete receptive vocabulary may be due to the age and subsequent life experience of older individuals with ID (Chapman, 2006; Facon et al., 2002).

There are also additional examples of relative conceptual vocabulary weakness in individuals with ID across different age groups. For example, younger children with Down

syndrome, fragile X syndrome, or fragile X syndrome with co-morbid autism have been shown to display lower conceptual receptive vocabulary ability than children who are typically developing and matched on non-verbal cognition (Price, Roberts, Vandergrift, Martin, 2007), but conceptual vocabulary appears to be commensurate with non-verbal cognition in adolescents with fragile X syndrome (Abbeduto et al., 2003). Additionally, although receptive vocabulary is considered a relative strength in school-age children with Williams Syndrome, there are significant differences in their performance on concrete versus abstract receptive vocabulary tasks (Mervis & Becerra, 2007). A similar pattern exists for individuals with ASD who have a relative strength in receptive concrete vocabulary when compared to syntax (expressive and receptive), and have receptive vocabulary that is significantly correlated with cognition (Kjelgaard & Tager-Flusberg, 2001). Receptive vocabulary ability in individuals with ID will likely impact their ability to understand different wh- question words and the words within the questions themselves.

Expressive vocabulary is also key in terms of answering questions. Many studies examining expressive vocabulary in individuals with ID have produced equivocal findings (Roberts, Price, et al., 2007). As in studies of receptive vocabulary, results are shown to vary according to etiology, age, and what is being compared. These findings can make it unclear whether this domain of language is a relative strength or weakness for individuals with ID.

For example, compared to expressive syntax abilities, expressive vocabulary has been shown to be a strength in children with Down syndrome (Laws & Bishop, 2003; Vicari, Caselli, & Tonucci, 2000). Additionally, expressive vocabulary has been shown to be commensurate with receptive vocabulary in children with fragile X syndrome (Roberts, Price, Barnes, et al., 2007), with non-verbal intelligence in children ID with mixed etiology

(van der Schuit, Segers, van Balkom, Verhoeven, 2011), and with non-verbal mental age in adolescents with Down syndrome (Laws & Bishop, 2003).

Expressive vocabulary has also been shown to be a weakness when compared to receptive vocabulary and cognition in other groups of individuals with ID. For example, it has been shown to be lower in children with Down syndrome and boys with fragile X syndrome and co-morbid autism spectrum disorders when compared to typically developing children matched on nonverbal mental age (Roberts, Price, Barnes, et al., 2007). These mixed results and considerable within-group variability (e.g., Roberts, Price, et al., 2007) make it difficult to parse out true strengths and weaknesses in expressive vocabulary in individuals with ID, but they do indicate that there may be weaknesses that could impact the ability to produce answers to some questions.

The majority of studies investigating receptive and expressive vocabulary in individuals with ID have used standardized assessments. Studies of language in connected speech samples (e.g., narratives, conversational language samples) have also produced important findings about the language used at the discourse level by individuals with ID. For example, children and adolescents with DS have been noted to produce fewer total and different words than typically developing peers matched on nonverbal mental age in both conversational and narrative language samples (Chapman et al., 1998). Studies of narrative language using wordless picture books have also revealed information pertaining to language directly related to the language children with ID could use to answer questions. For example, some analyses of narratives have looked at the use of "causal connectors" (e.g., "because," "so that") and "connectives" (e.g., "and") in production. Using these types of words could be important in answering "*why*" (causal language) and "*when*" questions (sequential language).

Adolescents and young adults with Down syndrome and fragile X syndrome have been shown not to differ from children who were typically developing and matched for nonverbal mental age in their use of causal connectors (Keller-Bell & Abbeduto, 2007). Additionally, adolescents with ASD who did not have a label of "high-functioning autism," children with ID of unknown etiology/learning, and typically developing children matched on mental age have shown no difference in the use of connectives (e.g., and, and then/so/but), adverbs and adverbial phrases (e.g., later), and causal connectives (e.g., because) during a narrative task (Tager-Flusberg & Sullivan, 1995). It should be noted that the ability to produce these words is still delayed in reference to age-matched peers who are typically developing.

Findings from studies of vocabulary in individuals with ID have several implications for understanding and answering wh- questions. First, it seems clear that although receptive vocabulary appears to be a relative strength across several populations of ID it continues to be an area of weakness relative to matched samples, and there seems to be a difference in the understanding of concrete and conceptual vocabulary. This could impact the ability to understand different conceptual levels of questions, and it points to possible differences in the acquisition of the ability to comprehend questions. Additionally, because expressive vocabulary seems to be more variable across different populations of individuals with ID there may be difficulty in producing specific, appropriate answers to questions. These aspects of semantic skills should be considered when investigating the understanding and answering of wh- questions in individuals with ID.

Syntactic Skills and Intellectual Disability

As with language skills in general, syntactic development in ID is generally thought to follow a similar but delayed course of development than what is seen in typical

development (Rosenberg & Abbeduto, 1993; Fowler, 1998). Similar to semantic skills, however, there are important distinctions that have been noted between individuals with ID and those who are developing typically or have language impairments. Because questions are examples of complex syntax, deficits in receptive syntax will likely lead to difficulty understanding questions. Likewise, difficulty with expressive syntax may impact the production of certain answers.

The influence of age has been investigated to determine if some populations of individuals with ID comprehend syntax better at different points of development. Indeed, age does appear to influence performance of some populations of individuals with ID (e.g., Roberts, Price, et al., 2007), but not others (Abbeduto et al., 2003; Chapman, 2006). For example, individuals with Down syndrome as early as age four are thought to have difficulty understanding simple sentences compared to children matched on mental age (Vicari et al., 2000). These differences are believed by some to become greater as individuals with Down syndrome get older, resulting in a distinct receptive syntax impairment (Laws & Bishop, 2003). It has even been postulated that this ability regresses as children with ID age (Abbeduto, Warren, & Conners, 2007; Chapman, 2006; Chapman, Hesketh, & Kistler, 2002). Interestingly, age also has appeared to play a role in individuals with ID of unknown etiology. In one study, individuals with ID and mental ages of 5 were found to have the same receptive syntax ability as a typically developing comparison group matched on mental age. In similar groups matched at mental ages of 7 and 9, there was a significant deficit in performance for the ID groups when compared to typically developing groups (Abbeduto, Furman, & Davies, 1989). Regardless of age, receptive syntax does not appear to be a relative strength in the way receptive vocabulary is for many individuals with ID, and it

appears to generally be lower than or commensurate with groups of individuals who are typically developing matched on mental age.

Across other populations and ages of individuals with ID, there are examples of how receptive syntax is related to cognitive and semantic ability. For example, adolescents and adults with fragile X syndrome have been shown to have receptive syntax skills that are synchronous with receptive semantic skills and nonverbal cognition (Abbeduto, Brady, & Kover, 2007; Paul et al., 1987). The same synchronicity has been found in adolescents and adults when the etiology of ID has been unknown (Chapman, 2006; Paul et al., 1987). However, in children with fragile X syndrome receptive syntax has been found to be below nonverbal cognitive expectations but not receptive vocabulary (Price et al., 2007). Additionally, receptive syntax has been shown to be closely linked to nonverbal cognitive ability in a sample of children with mixed etiology of ID (Facon et al., 2002), similar to receptive concrete vocabulary in individuals with Williams syndrome (Mervis & Becerra, 2007), but depressed relative to receptive vocabulary in individuals with ASD (Kjelgaard & Tager-Flusberg, 2001). These clear deficits in receptive syntax may be a factor in the comprehension of wh- questions in individuals with ID.

Similar variability in expressive syntax exists within and across populations of individuals with ID in the expressive syntax. Expressive syntax is also considered to be a relative weakness when compared to other domains of language and nonverbal cognition across various populations of ID (Fidler et al., 2007). Variability in this domain can be attributed to the same factors that cause variability in the other domains of language and it is logical to believe these types of weaknesses will likely impact the way answers to questions are expressed by individuals with ID.

Similar to other aspects of language, age is thought to impact the way individuals with ID use syntax to combine words. For example, in individuals with unknown etiology of ID, it is thought that the sequence of learning grammatical rules is generally similar to typically developing children who are matched for mental-age until the individuals with ID have an MLU of three. After this point, however, MLU levels are reached at later mental-age levels and shorter and less complex sentences are used (Paul, 2007). Studies of children with Down syndrome have shown that expressive syntax also progresses in the same order as typical development, but is more delayed than expected in comparison to nonverbal visual cognition (Chapman et al., 1998) and receptive syntax (Abbeduto et al., 2003). Unlike receptive syntax, however, expressive syntax has been shown to continue growing in this population through adolescence. Additionally, individuals with Down syndrome whose syntax is developing have been shown to use syntax that moves beyond the simple level in narrative contexts (Chapman et al., 2002).

Individuals with ID of other known and unknown etiologies also tend to struggle with expressive syntax, although, again, there is considerable inter-subject variability (e.g., Abbeduto, et al., 2007). For example, boys with fragile X syndrome have been shown to use shorter, less complex utterances in conversational language samples than boys who are typically developing and matched for mental age (Roberts, Hennon, et al., 2007). Others, however, have found expressive syntax to be at a similar level as the same type of control group (Abbeduto et al., 2001). Additionally, in a group of children with ASD who also have language impairment and could complete a standardized language assessment, expressive syntax was shown to be similar to receptive syntax in that they were both impaired relative to vocabulary (Kjelgaard & Tager-Flusberg, 2001). Finally, although expressive language is

considered to be a relative strength in individuals with Williams syndrome, it has historically been found to be synchronous with, rather than being spared in relation to, cognitive ability (Mervis & Becerra, 2007).

From these examples, it is clear that expressive syntax is a relative weakness for many populations of individuals who have ID. Weaknesses in this area will likely to lead to difficulty combining words in order to create cogent, detailed answers. This is especially the case in answering *how* and *why* questions. It may be the case that context plays a greater role in understanding questions and expressing answers for individuals with ID than it does for individuals who are typically developing because these questions and answers may be more easily understood if there is more context.

Questions and Individuals with Intellectual Disability

Understanding and answering wh- questions are complex tasks that require the integration of numerous language skills. Given the language difficulties of individuals with ID, it seems reasonable to assume that understanding and answering wh- questions will also be a source of difficulty. Question understanding and answering of wh- questions with a wide variety of different question words has not systematically been investigated in individuals with ID. This is remarkable given the role of wh- questions in supporting learning, assessing student knowledge, and promoting language development (Morgan, Moni, & Jobling, 2009).

There are, however, some studies that have examined how individuals with ID understand questions that vary in their syntax and conceptual level. Intervention studies designed to teach individuals with ID how to answer different wh- questions also provides valuable information regarding supports that may help individuals in this population understand and answer these questions. Additionally, wh- question word comprehension has been examined in individuals who have language impairment and intact comprehension. This

is important because it helps to elucidate issues that may be present in designing a study examining wh- question comprehension and answering in individuals with ID. As a whole, these studies inform what is known about how individuals with ID understand questions

Comprehension of Wh- Question Words in Individuals with Language Impairment

Individuals with language impairments demonstrate age-related differences in their ability to understand and respond to questions. Parnell, Amerman, and Harting (1986) looked specifically at the ability of individuals with language impairment aged 3 to 7 to comprehend questions in different referential conditions. The children responded to a battery of wh-questions that had been used in a prior study with age matched participants without disabilities (Parnell et al., 1984). The results were then compared across the two studies.

In comparing the two groups, Parnell and colleagues (1986) found age-related differences in the number of "functionally appropriate" and "functionally accurate" responses participants provided. That is, the older children responded with more of both types of answers. Furthermore, the children with language impairments were less successful answering questions when a referent was not present, and younger children had more difficulty with these questions than older children. Additionally, the lack of an immediate referent led to differences in terms of the number of functionally appropriate responses children provided. That is, when a referent was not present, children had more difficulty providing an answer related to the category. The only question word that was more difficult for the children with a referent present was *when*. In this case, the participants were able to answer questions without an immediate referent better than those about an object. However, the researchers did not report how well the children answered *when* questions about the picture.

In terms of the actual question forms, *why*, *when*, and *what*- happened were more challenging for the children to answer correctly than which, *what* + be, *where*, and *who*se. There were general differences between the students with language impairments and those who were typically developing in terms of functional appropriateness and accuracy. This difference was most extreme in the answers provided to *when* and *why* questions. The results of this study suggest that students with language impairments, like those who are typically developing, may answer questions in functionally appropriate ways before they do so accurately.

Lee and Ashmore (1983) also investigated receptive understanding of a variety of wh-questions. They studied 20 children ranging in age from 4;3 to 6;4 who met their criteria of having a delay in either receptive language, expressive language, or both. Four different *who*, *what*-do, *where*, *when*, *why*, and *how* questions respectively, were asked about the same storybook used in that study. The study involved determining the relative ease of understanding each question type. and the patterns of errors participants made. The categories used for the errors included: substitutions (e.g., answering a *why* question like a *where* question), a patterned response (e.g., "five minutes" for *when* or "because" for *why*), repetition of a portion of the question, no response, and a completely inappropriate response.

Consistent with the literature on typical development, children had success answering questions correctly in the following order (from easiest to most difficult): *where*, *what*-do, *who*, *why*, *when*, and *how* questions. In their error analysis they found that the error strategies used in incorrect answers to *why*, *when*, and *how* developed in stages. For *when* and *why*, the children typically used some type of substitution, followed by a patterned response, and finally produced correct responses. In response to *how* questions, children provided

substitution responses followed by correct responses. No patterned responses were given. Both of these studies indicate that children with language impairments, although delayed, seem to follow the same progression in terms of question answering as children who are typically developing.

Comprehension of Wh- Questions in Intellectual Disability

One study has specifically investigated the ability of individuals with ID to answer different types of wh- questions; however, this study was not designed to explore the influence of wh- words on comprehension, rather it looked specifically at the influence of syntax on comprehension. Joffe and Varlokosta (2007) investigated whether there were differences in the way children with Down syndrome, Williams syndrome, and typically developing children answered *who* subject, *what* object, and subject and object which-noun phrase questions. The groups of children with Down syndrome and Williams syndrome were matched on performance IQ, chronological age (8;7 and 8;9 respectively), and mental age (4;6 and 4;8 respectively). The typically developing group was matched on mental age to the groups of children with Down syndrome and Williams syndrome. Participants were asked questions about a story that was acted out. No significant main effect was found for comprehension of questions based on type (e.g., subject vs. object) across the groups of participants. Additionally, the group of children with Williams syndrome performed better than the group of children with Down syndrome on the comprehension of all questions when they were aggregated. However, the performance of both the Down syndrome and Williams syndrome groups was still only at 57% and 43% receptively (33% was chance level), whereas the typically developing group answered questions correctly 83% of the time.

Intervention Studies and Comprehension of Wh- Questions

Although there has been a lack of studies investigating comprehension of different wh- question words in individuals with ID, there are a variety of intervention studies that have been conducted with individuals with autism and co-morbid ID. These are important because different methods of presentation were used to facilitate understanding of these wh-question words. These intervention studies informed the method of presentation employed in the current investigation. For the purposes of this review, results with children who appear to have co-morbid ID based on participant description will be highlighted.

Krantz, Zalenski, Hall, Femske, and McClannahan (1981) were the first to provide a question-answering intervention that targeted the question words *what*, *why*, and *how*. In this multiple-baseline design across the question words, two students were taught to answer questions about a magazine picture. The participants were successful in learning to respond in a complete sentence using responses that have been described as categorically related and accurate in research involving children without disabilities. Both students reached criteria for the wh- question words, which required them to answer the questions correctly at least 80% of the time on untaught probes.

Secan, Egil, and Tilley (1989) sought to replicate the findings of Krantz et al (1981). They engaged four children with ASD (ages 7;11 to 9;2) studying an intervention targeting responses to *why*, *how*, and *what* questions across different contexts (i.e., storybook, natural-context). All students were able to answer questions during the magazine training trials, but were less successful with the generalization tasks. With booster training sessions, all students but one increased generalization in the storybook context. Importantly, the authors examined how well the children answered questions when there was a referent available in the picture

cue (i.e., all *what* questions, *how* means and action questions) and those that could not be answered by referring to something in the picture (i.e., all *why* questions, *how* affect questions). Little difference was found in the number of training trials to criterion, but there was a difference in the generalization probes with children answering 83% of the questions with and 42% of the questions without a visual cue. Additionally, generalization was higher for *what* as opposed to *how* and *why* questions. The fact that these individuals had difficulty with more conceptually difficult questions and benefitted from visual support in answering some questions influenced the methods and research questions in the current investigation.

While picture referents clearly influence question comprehension for children with ID, there is also evidence that they can learn to answer questions in the absence of picture clues. Jahr (2001) taught five participants with ASD (ages 3;11 to 7;2) and "mild" to "moderate" ID to answer questions without visual cues. Using broad *what*, *where*, *who*, and *why* questions with answer frames that matched the syntax of the question as closely as possible (e.g., the answer to the question "*What* do you like to drink?" would be "I like to drink ____."), Jahr taught the students to respond in complete sentences. Maintenance probes taken 4 to 7 months later indicated that the participants were generally able to maintain their ability to answer these types of questions.

Hundert and Delft (2009) conducted the most recent study investigating ways to teach individuals with ASD and ID how to answer wh- questions. Unlike the previous studies, they sought to teach students only how to answer inferential *why* questions. Only one of the three participants in this study had intellectual disabilities (IQ of 70 and an adaptive behavior composite of 57) so only his results are discussed here.

This intervention consisted of asking inferential *why* questions that were either based on general information, a verbal story, or a sequence of picture cards. The target participant achieved mastery across all conditions, but was able to answer the questions about the picture sequence more quickly than those with a verbal story. Furthermore, he answered questions about verbal stories with decreasing accuracy in generalization probes. Interestingly, he (as well as the other participants) was unable to answer the *why* questions in the alternate formats (i.e., questions about general information and sequence cards) until he was trained to do so. These findings influenced the decision to compare question comprehension relative to a single picture and a personally relevant general routine in the current investigation.

Comprehension of Different Types of Questions in Intellectual Disability

Rather than strictly classifying questions based on the type of wh- word used, comprehension of different types of questions has been investigated in individuals with ID. For example, some researchers have investigated literal and inferential question comprehension (Zetlin & Gallimore, 1983; Hewitt, 1998). Additionally, these studies have occurred across different contexts. Information from these studies is important for understanding how individuals with ID may comprehend wh- questions given various conditions and formats.

Zetlin and Gallimore (1980; 1983) examined the ability of individuals with ID to learn to comprehend questions with different conceptual levels. In this study, three students ranging in age from 12;0 to 14;9 years old took part in an intervention designed to increase their listening comprehension through a questioning technique. This technique was intended to encourage the use of higher-order, self-regulatory strategies. All students had IQs that ranged from 40 to 50. The intervention occurred 3 times per week for a total of 23 sessions.

The technique used in the study was coined "responsive questioning" as questions were continually adapted to meet the student's needs while listening to basal reading texts, moving from lower-order, concrete questions to higher-order, more abstract questions.

The students were able to answer higher-level questions with appropriate scaffolding and questioning that was adapted based on the student's initial responses. Student's showed that they were able to draw inferences when listening to these pre-primer level stories and answer both higher and lower conceptual level questions with appropriate support. This study informed the decision to categorize the wh-questions in the current study into more concrete and abstract categories.

Hewitt (1998) also examined how individuals with ID answer different types of questions. In this study, a random selection of transcripts was analyzed from weekly sessions during which adolescents with ASD presented information to the rest of the group about school and individual activities. All students were labeled as being in the mild-to-borderline range of intellectual functioning. The researcher posed questions to the group about the information they shared. The types of questions that were analyzed were: questions that were longer than 7 words, questions with multi-clausal syntax, questions requiring inference, and indirect questions (e.g., questions that appear to be yes/no questions but actually require the listener to provide more in-depth information). Responses were categorized as either being adequate or inadequate.

There were no statistically significant differences between the inadequate responses of any of the categories. However, some trends emerged when looking at the percentage of inadequate responses for the categories. The participants provided adequate responses when shorter sentences that had simple syntax were asked of them. However, the participants still

had difficulty with shorter questions that required them to make inferences. These results informed the decisions in the current study to keep questions as syntactically simple as possible in order to focus on the impact of the specific wh- question words and to consider the inferential load of questions.

Summary of Research Regarding Questions and Individuals with ID

Although a majority of the studies of question comprehension involving individuals with ID did not directly investigate the comprehension of questions with different wh- words, they all provide information that contributes to understanding question comprehension in this population. For example, these studies indicate that there may be differences in the way individuals with ID comprehend questions of different conceptual levels (Hewitt, 1998; Zetlin & Gallimore, 1983). Additionally, these studies provide valuable information about the impact of syntax (Hewitt, 1998; Joffe and Varlokosta, 2007) and the way questions are presented (e.g., Hundert and Delft, 2009) on the comprehension of questions of individuals with ID. All of this information informs the way questions should be measured in this population and confirms the importance of studying this aspect of language.

Conclusion

Understanding and answering questions are complex skills that involve the interaction of different language and cognitive skills. There is a relatively clear developmental sequence of the comprehension of different wh- question words in typical development (e.g., Tyack & Ingram, 1977). In individuals with language impairment, the developmental sequence of wh-question comprehension appears to be the same albeit a little slower (Lee & Ashmore, 1983). Additionally, the referential source of wh- questions has been shown to impact

comprehension of wh- questions in both children who are typically developing and those with language impairments (Parnell et al., 1984; Parnell et al., 1986).

The comprehension of wh- question words has not been examined in individuals with ID. Given the difficulties that individuals with ID are likely to have across receptive and expressive domains of language, it is probable that they will have difficulty understanding and answering questions. This study is an attempt to investigate how students with ID answer different wh- questions that are presented in two referential conditions. Additionally, questions will be grouped and compared based on their conceptual complexity. Further, an analysis of errors will be conducted to determine whether there are distinct patterns that emerge when individuals with ID did not understand questions. The results of this study will add to what is currently known about question answering in this population.

CHAPTER 3

Methods

The purpose of this study was to investigate the influence of wh- question words, their relative conceptual level, and a picture referent on the comprehension of and answers to questions in individuals with intellectual disabilities (ID). A secondary purpose was investigating the types of errors individuals make when they provide an incorrect response. Given that understanding wh- questions influences communication, literacy, and academic development, clarifying our understanding of the relative difficulty of different types of wh-question words used with and without picture referents may ultimately influence the success individuals with ID experience across these domains.

Research Questions

Using a combination of researcher-designed and standardized instruments, this study investigated the comprehension of questions with different wh- question words with and without a picture referent in school-aged individuals with ID. This was accomplished by asking the participants to respond to questions under two conditions (i.e., with and without a picture referent). The specific research questions were:

1. What wh- question words (i.e., *who, what, where, when, why, how*) do school-age individuals with ID comprehend most successfully?
    a. Is there a difference in wh- question word comprehension between more concrete (i.e., *who, what, where*) and abstract (i.e., *when, why, how*) question

b. words across two question-answering conditions in school age individuals with ID?

c. Does the condition in which the questions are presented have an effect on the relative ease of answering these questions for school-age individuals with ID?

2. What is the relationship between receptive vocabulary and question comprehension in school age individuals with ID?

    a. Is there a relationship between overall wh- question comprehension and receptive vocabulary?

    b. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary ability based on whether the question words are more concrete or abstract?

    c. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary based on whether the questions are asked in the picture or no-picture condition?

    d. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary ability based on whether the question words are more concrete or abstract within the picture and no-picture condition?

3. What types of responses do school-age individuals with ID give when their response to a question is incorrect?

    a. Do patterns of errors exist in the incorrect responses school-age individuals with ID provide (i.e., categorical relatedness, substitutions, I don't know/no responses, topic related, unrelated)?

b. Do errors or patterns of errors in the incorrect responses of school-age individuals with ID vary based on the use of a picture referent?

Participants and Setting

Participants

Thirty-nine students were included in this study (22 male and 17 female). All participants were concurrently enrolled in a larger investigation of a yearlong literacy intervention for students with ID. All participants in the larger study who could provide a spoken response were included in the initial group that was given the current assessment. Per the requirements of the larger study, all participants were between the ages of 8 and 19. The average age of the students was 13 years 11 months (standard deviation 2;03, range 8;08 - 19;05). Their grade placements ranged from 3rd to 12th grade with 5 students enrolled in grades 3 to 5 (2, 2, and 1 respectively), 28 enrolled in grades 6 to 8 (9, 8, and 11 respectively), and 6 enrolled in grades 9 through 12 (0, 1, 1, 3, and 1 respectively). Two of the students were Hispanic or Latino, 24 were African American, 10 were white, and 3 had mixed ethnicity.

All of the students had some level of intellectual disability as measured by the school system on a variety of standardized measures used as part of their mandated identification and evaluation process. The primary exceptionalities for the students included Intellectual Disability (n = 25), Multiple Disability (n=5), Autism (n=8), and Other Health Impaired (n=1). Thirty-eight of the 39 students included in the study were educated in separate special education classrooms for more than 60% of their educational time. Additionally, 61% of the students had a reading or speech-language goal on their IEP related to either the comprehension or asking of questions.

Setting

The participants were drawn from three school systems across North Carolina. The first system is located in central North Carolina. As a whole, the school district is ethnically diverse with approximately 52% of its population being African American, 21% white, and 21% Hispanic. There are approximately 33,500 students in this district. The second school system is located in the north central part of the state. It is ethnically diverse as well as with 53% of its 17,500 students being African American, 7% Hispanic, and 36% white. The final district is smaller than the other two with a total of 7,100 students and is located in central North Carolina. This school district is less ethnically diverse than the other two districts with approximately 17% of the students being African American and 11% Hispanic.

Study Assessment Measures

Question Comprehension Battery.

A measure was created by the primary researcher to probe comprehension of the semantic meaning of questions that use particular wh- question words (i.e., *who*, *what*, *where*, *when*, *why*, and *how*) across two different conditions (i.e., with and without pictures). The question comprehension battery and stimuli, as well as the entire study protocol, was approved by the Institutional Review Board (IRB) at the University of North Carolina at Chapel Hill.

Six different question words were assessed: who, what, where, when, why, and how. These question words were chosen for inclusion in this study because they have been examined in other studies of wh- question word comprehension (e.g., Ervin-Tripp, 1970; Tyack & Ingram, 1977) and are used widely in academic settings (see Common Core State Standards, 2010). Many of the wh- words themselves can have different meanings (i.e., how, what, when, why). In the current study, the who question required participants to identify a person, the what question asked students to identify items, and the where question required

identification of a place. The when question addressed time identification, the why question required the identification of a cause of an action, and the how question was procedural. Using these forms allowed the questions to be divided into two categories based on their relative concreteness versus abstractness and developmental order (Bloom et al., 1982; Owens, 2008). The who, what, and where questions were classified as concrete because they refer to things that can be seen; that is, they reflect a concrete person or location. The when, why, and how questions were classified as abstract because they rely on a different level of conceptual ability. In this case, the concept of time is needed for when, causality for why, and the ability to relate a response to a procedure is necessary for how (Owens, 2008).

The questions were all about lunch. Lunch was selected because it is a common experience among all participants across the school systems. They all eat lunch at school and all of the schools have similar cafeterias and lunch schedules. The decision was made to select a common, repeated experience as a topic to reduce the number of other abilities that would be tapped (e.g., working memory or world knowledge). Using a more traditional, de-contextualized approach to questioning such as reading or telling a story would have tapped these other abilities to an extent that would introduce too much variability in this initial investigation.

Picture Referent Condition. The conditions provided different levels of support for the students in terms of the presence of a picture referent. Many studies that have examined question answering in children have employed similar procedures, asking participants to respond to questions with a picture available (e.g., Ervin-Tripp, 1970; Parnell et al., 1984; Tyack & Ingram, 1977). In this study, a single picture was chosen to reduce the variability inherent in introducing different pictures for different question words. For example, answers

to questions posed with different pictures may be differentially influenced by aspects of that picture (e.g., knowledge of the people in the picture, the setting, actions). The picture used (see Figure 3.1) depicted a cafeteria scene that should have been familiar to all of the students based on their school experience. This picture provided students with a clear referent for each question they were asked. Who, what, and where questions could be answered labeling the correct parts of the picture. The picture also provided additional context for answering the when, why, and how questions. Although a direct referent was provided in the picture condition, these types of more abstract questions required the students to make an inference. In the second condition, no picture referent was provided.

Morphosyntax and Vocabulary of the Questions. The questions that were used in the investigation are provided in Table 3.1. Because of the difficulties individuals with ID have with receptive syntax and vocabulary (e.g., Paul, 2007), both morphosyntax and vocabulary were controlled in the construction of the questions. Previous studies of question comprehension have used "syntactic frames" (e.g., Where is the boy (subject) eating (verb – ing)?) using the same core vocabulary matched with different question words (e.g., Tyack & Ingram, 1977). Using the same frame for all questions, however, causes difficulty when certain question words are matched with stimuli. For example, asking questions about characters in a picture requires slightly different syntax and vocabulary than asking a participant about his or her own life. If an exact syntactic frame was employed, the resulting questions would have been bizarre or contextually inappropriate (Parnell et al., 1984). Since a frame could not be employed for these reasons, all questions were created so they were close in length (e.g., by employing word and morpheme counts) and had as simple morphosyntax as possible. Each word was one morpheme in length in each condition. The

questions in the picture condition were 6 words in length, while the no-picture condition questions were 5 words in length. The only difference between the two conditions was that in the picture condition the article "the" was used to ask about "the boy," while the questions in the no-picture condition referred to "you." Additionally, grammatical categories of words (e.g., pronouns, prepositions, verb tense) were selected to include only categories that are expected to be understood at a 48-month developmental level or lower.

Table 3.1
Researcher developed battery of questions

Picture	No Picture
Who does the boy eat with?	*Who* do you eat with?
What food does the boy like?	*What* food do you like?
Where does the boy eat lunch?	*Where* do you eat lunch?
When does the boy eat lunch?	*When* do you eat lunch?
Why does the boy eat lunch?	*Why* do you eat lunch?
How does the boy get lunch?	*How* do you get lunch?

All questions used in the study were object questions. Although some studies have reported that subject questions are better understood by younger children, this point is equivocal (e.g., Stromswold, 1995) and has varied across different question words in individuals with language impairment (Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2011). Further, with the wh- question words that were being assessed, subject questions could only have been made with *who* and *what*. For these reasons, object questions were selected to ensure that the questions were as parallel as possible across all question types and conditions.

All questions used in the study were object questions. Although some studies have reported that subject questions are better understood by younger children, this point is

equivocal (e.g., Stromswold, 1995) and has varied across different question words in individuals with language impairment (Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2011). Further, with the wh- question words that were being assessed, subject questions could only have been made with *who* and *what*. For these reasons, object questions were selected to ensure that the questions were as parallel as possible across all question types and conditions.

Figure 3.1
Picture Referent Used with Questions

In addition to controlling for morphosyntax, the frequency of vocabulary used in the questions was controlled across conditions. Because there is no corpus of words that lists the

frequency or dispersion of words in spoken American English, the frequency and disperson of words in written English served as a proxy. The Educator's Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995) served as a guide to insure that all of the words employed across questions and conditions occur at extremely high frequencies and dispersion across contexts (reported as a U-statistic in the guide). All words that were used in the questions have a U-score above 1.0 on the first grade list in the guide. U-Scores for each word used in the stimuli are provided in Table 3.2.

The Peabody Picture Vocabulary Test – 4th Edition

The Peabody Picture Vocabulary Test – 4th Edition (PPVT-4; Dunn & Dunn, 2007) is a standardized instrument used to assess one-word receptive vocabulary level. This test requires students to point to a picture of a target word from a field of four after it is presented orally. This measure was administered to all students participating in the research study. This test is appropriate for individuals from the ages of 2;6 to 81. It has been used in several studies of individuals with ID. The PPVT-4 has internal consistency reliability for participants from the ages of 8 to 17 of .93-.95 (coefficient alpha), .86-.95 (split-half), and .93-.94 (test-retest). The PPVT-4 is a measurement of a single domain of language, single-word receptive vocabulary. Given that one purpose of this study was to examine the comprehension of semantic meanings of question words, the PPVT-4 was an appropriate measure to compare the overall relationship of receptive vocabulary to the comprehension of these specific semantic targets.

Procedures

Testing occurred 1:1 in quiet settings at the schools where participants were enrolled over the course of approximately four weeks in the spring of 2011. All students completed the PPVT-4 (Dunn & Dunn, 2007) first and then completed the question-answering battery.

There were a few students who completed the question-answering battery at a separate time from the PPVT-4.

Table 3.2
U-Scores for All Words in All Pictures

Words	Overall U Score	First Grade U Score
boy	292	292
do	2102	4560
does	541	552
eat	270	1051
food	570	454
how	1616	2048
like	1810	3487
lunch	66	407
the	68006	50950
what	2560	4928
where	1073	1477
who	1826	1086
why	597	1218
with	5844	4181
you	7600	15472

Prior to the presentation of questions in both conditions, the assessor provided the students with the following short set of instructions:

- Picture condition: "I'm going to ask you some questions about a picture. In this picture there is a boy and a girl eating lunch (point to each person in the picture). I want you to listen carefully. I want you to do your best to tell me the answer to the questions I ask. Ready?"
- No picture condition: "I'm going to ask you some questions about lunch. I want you to listen carefully. I want you to do your best to tell me the answer to the questions I ask. Ready?"

In both conditions, repetitions were allowed if the student indicated that they did not hear the question or if there was a reason a repetition was necessary (e.g., student distraction, interruption of the task). Assessors were instructed not to prompt the students in any way. Assessors were allowed to ask for clarification or repetition if they did not understand the student's response.

Assessor Training

Five assessors who worked on the larger study were trained as a group in the administration of the question-answering battery. One assessor had a Ph.D. in education and two others had Ph.D.s in the speech and hearing sciences. A doctoral student in education and a masters-degree student in the speech and hearing sciences also administered the battery. Each assessor was provided with a file with randomly ordered testing protocols and a picture to be used in the picture condition. All rules regarding the administration of the battery were reviewed until the primary researcher determined that the assessors were reliable test administrators.

Assessors were given pre-printed, numbered forms that presented the questions in random order as generated by a random number generator. Assessors were instructed to alternate between picture-first and no-picture-first forms each time they administered the question-answering battery. The process resulted in 20 students who completed the picture condition first and 19 who completed the no-picture condition first. All questions were presented in random order across participants

Recording of Responses

Assessors were asked to record the responses of the participants in two ways. First, they were instructed to write down everything the students said in response to the question on

the test-recording forms. They were also asked to write notes of their own reflections of the students' responses (e.g., whether a response was difficult to understand, the direction a student may have pointed). Additionally, they audio-recorded the responses of students whose caregivers had given consent. Of the 39 students included in the sample, 28 were recorded. Seven students did not have permission from their parent/guardian to be recorded. Additionally, four were not recorded because of a technical error with the recording equipment.

Response Inclusion Criteria

All participants from the larger study who could provide intelligible verbal responses to the questions used in the current study were assessed. Students with uncorrected vision impairments were not included in this study. Of the 84 students who participated in the larger study, 50 were originally selected by the trained assessors to be complete the question battery. Participants who scored within two standard deviations of the mean on the PPVT-4 (Dunn & Dunn, 2007) were removed from the analysis. Given that approximately 95% of individuals taking the PPVT-4 nationally fall within 2 standard deviations of the mean (standard score higher than 70), selecting this cut-score added one more indicator that participants fell into the 2.5% of the population with ID. The cutoff point was also chosen to minimize possible ceiling effects on the question-answering battery. Six participants were eventually excluded from the final data set because their PPVT-4 standard score was within two standard deviations of the mean. Of the remaining 44 students, thirty-two were audio-recorded. These audio recordings were subjected to an intelligibility reliability process (described below) to determine whether they could be included in the sample or not. At least

8 of the student's 12 responses had to be intelligible for the student's responses to be included in the study.

Besides intelligibility, there were other criteria necessary for participants' responses to be included in the analyses. For example, some of the analyzed answers that students provided were only partially intelligible. For instance, in response to the question "Where do you eat lunch?" one student replied with a response that was unintelligible at the beginning. He then clearly said "in the cafeteria." If enough of the response could be understood to determine whether the student provided a reasonable response to the question or not, it was included in the analysis. Therefore, this student's response was included in the analysis. There were three total responses included in the final corpus of analyzable responses that were like this. In addition, two of the eleven students whose sessions were not audio-recorded had responses that could not be understood by the assessor who worked with them. These assessors indicated on the protocols that the responses were difficult to understand. This resulted in the elimination of three responses for each of the two students.

Other factors also impacted inclusion for analysis. For example, some students did not respond to questions with a verbal response. As no-responses are a legitimate indication of non-comprehension in students who can speak, these were included in the analysis as an incorrect answer as long as eight of the twelve questions were answered verbally. Additionally, if a student pointed to the picture or a place in the testing environment that was observed and noted by the assessor it was counted as the answer if the referent was clear (e.g., if the student pointed to the "table" in response to a question). If it was not clear, this item was not included in the analysis. In this study, four responses were excluded based on

these no-response and pointing criteria. Only one student was removed from the sample because he had five "no responses."

Transcript Cleaning

The primary researcher transcribed all responses that were audio-taped. Each intelligible word was transcribed verbatim and a special code was used to mark a word or words that were difficult to understand. When possible, the primary researcher wrote what he believed the words were on these parts that were difficult to understand. A second trained research assistant listened to the audio-recordings while reading the transcripts and marked any discrepancies.

The primary researcher and research assistant then met to come to consensus on discrepancies in the transcripts. There was a very high rate of agreement and very few discrepancies. Of the 33 audio recordings that were transcribed, five were excluded because there were more than four responses that both researchers agreed were unintelligible, were no-responses, or had a pointing response with no clear visual or verbal referent. Of the 28 remaining audio-recorded transcripts, 13 items were removed from the final analysis due to the exclusion criteria described above.

The primary researcher then counted all of the words from the audio-taped transcripts in the responses that were included in the final data set. The initial disagreements between the primary researcher and secondary listener were then counted on a word-by-word basis. They initially disagreed on five words. Four of these were in the same answer. They agreed on 840 words of the 845-word sample (99.4%). Consensus was obtained on those remaining five words.

Scoring the Responses

The primary researcher developed a coding scheme and accompanying manual (see Appendix A) designed to judge the answers as being correct or incorrect. Correct responses for more concrete questions (i.e., *who*, *what*, *where*) were easier to define than correct responses for the more abstract questions (i.e., *when*, *why*, *how*). This is not surprising given that the responses to more abstract types of questions are less constrained than those that have answers that can typically be answered with nouns or prepositional phrases (i.e., the more concrete questions). The *why* and *how* questions could be answered in multiple ways. The primary researcher used the written guidelines in the manual to score the responses as either correct or incorrect.

A second scorer completed reliability scoring on all responses. This scorer holds a Ph.D. in education and has experience working with students who have ID. Prior to this reliability scoring process, a meeting was held in which the scoring scheme was discussed. Additionally, the second scorer used the coding manual. Point-to-point inter-rater reliability was conducted by dividing the number of agreements by disagreements for all answers to all question words in each condition. Table 3.2 shows the inter-rater reliability for the primary researcher and the second scorer. To resolve disagreements, the primary researcher considered the comments of the second scorer, consulted the manual and rescored.

Table 3.3
Percentage of Agreement of Correct/Incorrect Answers per Question Word

	Who	What	Where	When	Why	How
Picture (% agreement)	97.3%	91.9%	94.7%	94.9%	94.7%	100%
No Picture % agreement)	97.3%	94.6%	100%	89.5%	97.3%	100%

Error analysis

To detect and describe patterns in the incorrect answers of the questions, six different codes were developed by the primary researcher. Similar to the development of the correct/incorrect codes, these were described and compiled into a manual (see Appendix B). Previous research has noted developmental differences and differences between individuals who are language impaired and typically developing in terms of "functionally appropriate" (those that are categorically related) and "functionally accurate" (those that are right) answers to questions (Parnell et al., 1984). Further, these types of distinctions have been shown to be more apparent in children with language impairment (Parnell et al., 1986). The first error code was used to code all incorrect responses that included elements appropriate to that category of question. For example, if an incorrect response to a *why* question contained a causal word, the item was scored as being "categorically related."

The second error code was used for errors that resulted from a "substitution" process. Specifically, if an incorrect answer could have been answered correctly with the substitution of an alternate question word it was coded as a substitution (e.g., answering "*When* do you eat lunch?" with "in the cafeteria"). To be given credit for a substitution, the answer needed to be correct for the question it was substituted for. For example if a student answered the *why* question with a *when* response, but the time was 5:30, it would not be considered a substitution because lunch is not eaten at 5:30. The *where*, *when*, *why*, and *how* questions had clear substitutions that could have been made based on the syntactic frames of the questions. The *who* (e.g., "Who do you eat with?") and *what* (e.g., "What food do you like?") questions had fewer clear substitutions with other wh- questions. For example, no question word could

be substituted for *what* in "*What* food do you like?" *What* could have been substituted for *who* in "Who do you eat with?"

The next error code was designed to classify answers that were scored as wrong because the student did not offer a response (i.e., "no response") or said "I don't know." The fourth error code was used for incorrect responses that were partial or complete repetitions of the question and only included words that were used in the question. Finally, the remaining responses were examined to determine their relatedness to the questions that were asked. Incorrect responses were considered to be topic related if they had something to do with lunch (e.g., "soda") or about things in the picture (e.g., "chairs"). The remaining answers were classified as being unrelated.

Inter-rater reliability followed the same process and was conducted with the same second scorer. Point-to-point agreement was determined by dividing the number of agreements by disagreements of the incorrect responses for all coding categories within each question word and condition. Agreement was calculated to be 92.7% for all codes, question words, and conditions combined.

Data Analysis

Descriptive statistics, logistic regression, and error analysis were used to address the research questions.

Descriptive Statistics

Descriptive statistics were initially used to answer the first research question regarding differences between the answering of wh- question words. Additionally, differences in the percentages of questions that were answered correctly were described in terms of whether they were asked in the picture or no-picture condition. Success in answering these questions was described relative to each of these domains.

Logistic Regression

A 6 x 2 repeated measures logistic regression model with multiple observations was constructed to determine whether the probability of answering questions with one type of question word correct was statistically different from the probability of answering a question with a different question word (e.g., *how* vs. *when*). A 2 x 2 repeated measures logistic regression analysis was then run with question words grouped together based on conceptual level (i.e., *who, what, where* vs. *when, how, why*). Together these analyses allowed the testing of research questions regarding differences in answering different wh- questions, conceptual groups of questions, and the impact of condition on the probability of answering those questions correctly. The individual contrasts were completed using IBM SPPS (19.0). The difference between groups of questions was analyzed using SAS (9.2) statistical software.

To address the second research question, the same logistic regression models were utilized. However, the PPVT-4 (Dunn & Dunn, 2007) raw scores were entered into the model as a covariate in order to analyze its impact on the probability of answering the questions correctly. The PPVT-4 raw score was chosen because age occasionally resulted in floor effects using standard scores for older students even though they answered more items correctly than younger students.

The first model analyzed the impact of receptive vocabulary on overall performance on all of the question words combined. Then the analysis was repeated to investigate the impact of receptive vocabulary on the two conceptual categories of questions (*who, what, where* and *when, why, how*). The second model analyzed the impact of PPVT-4 (Dunn & Dunn, 2007) raw scores on the individual contrasts of the question words (e.g., *when* vs.

what). Further, the influence of the PPVT-4 raw score on the probability of being able to answer the individual contrasts, conceptual groups, and overall questions across conditions was analyzed. The influence of the PPVT-4 on the overall number of questions answered correctly and the wh- question word contrasts was analyzed using IBM SPPS 19.0. The influence of the PPVT-4 scores on the difference between groups of questions was analyzed using SAS 9.2 statistical software.

The final component of the second research question investigated whether there was differential relationship between receptive vocabulary ability and the probability of answering questions correctly based on the conceptual level of the questions. This relationship was investigated in the picture or no-picture condition. A logistic regression model was again created in which the PPVT-4 (Dunn & Dunn, 2007) raw scores were used as a covariate. Similar to previous models, estimates of the probabilities were created based on the categories of question words. In this model, however, the estimated probabilities were compared based upon whether students fell below, within, or above a standard deviation of the samples raw scores on the PPVT-4. Models were created for all of the questions words combined and within each condition.

Analysis of Incorrect Responses

All responses were analyzed using the error analysis coding scheme. All transcriptions of incorrect responses were entered into a Microsoft Excel 2007 spreadsheet. Descriptive statistics were calculated for each category, within each condition, and subsequently described.

Summary

The current study utilized a combination of repeated measures logistic regression and descriptive statistics to determine which wh-questions students with ID understood most

successfully, if the probability of answering these questions correctly was related to their receptive vocabulary ability, and if this varied across conditions. Additionally, answers were analyzed to determine if any patterns existed across incorrect responses. To determine the answers to these questions, a question comprehension battery was carefully designed and administered to the participants. Additionally, strict inclusion criteria were developed to determine what student answers could be included in the analysis. Further, reliable codes for incorrect/correct responses and error analysis were developed. These were used as the basis to address a total of 10 research questions

CHAPTER 4

Results

The primary purpose of this study was to assess the comprehension of wh- questions with and without a picture referent in individuals with intellectual disabilities (ID). Secondary purposes included investigating the relationship between single word receptive vocabulary and question comprehension and describing patterns of incorrect responses to questions. Answers were judged to either be correct or incorrect following a scoring guide developed by the primary researcher. Incorrect answers were examined for their categorical relatedness to the particular wh- question word that was assessed, whether an "I don't know/No response" was provided, and whether the response was a repetition of part of the question. Further examination of these incorrect responses was conducted to determine if they would have been an acceptable response if a different wh- question word was substituted for the one in the target question. The remaining responses were evaluated to determine if they were related to the topic, lunch, or something that was in the picture referent.

Descriptive statistics were interpreted in order to draw conclusions about the comprehension of questions with different wh- question words, the condition in which they were asked, and their conceptual level. Additionally, logistic regression models with multiple observations between subjects were constructed in order to evaluate relationships between success in responding to questions with different wh- question words and the participants' one-word receptive vocabulary ability. Analyses were conducted in both IBM SPSS (19) for

Windows and SAS (9.2). All analyses were interpreted with the alpha set at .05. Raw scores were used in all analyses conducted with the PPVT- 4 (Dunn & Dunn, 2007). Descriptive statistics were interpreted and described in the error analysis. All results are described in reference to the study's research questions.

Comprehension of Wh- Question Words

The first research question investigated the relative ease of answering the question types across conditions. Specifically, research question 1 with its two sub-questions was:

What wh- question words (i.e., *who, what, where, when, why, how*) do school-age individuals with ID comprehend most successfully?

a. Is there a difference in wh- question word comprehension between more concrete (i.e., *who, what, where*) and abstract (i.e., *when, why, how*) question words across all question answering conditions in school age individuals with ID?
b. Does the condition in which the questions are presented have an effect on the relative ease of answering these questions for school-age individuals with ID?

To answer these questions, each participant response was coded as correct or incorrect following the coding rules established by the primary researcher. Table 4.1 displays the descriptive statistics for each of the question words in each condition. These percentages are the average number of responses that were correct for questions with each question word. Additionally, questions with different question words are combined into groups based on their conceptual level (e.g., more concrete versus abstract answers).

Investigation of the descriptive statistics indicated that the mean percentages for all question words across conditions (e.g., *who* picture vs. *who* no picture) were similar. Only,

what, *why*, and *when* were at least 10 percentage points different from one another across conditions. Additionally, only the *why* questions appeared to be answered more successfully in the picture condition. The difference between conditions for the *when* question is larger than any other question word (20%). Grouping the comprehension of questions based on condition and conceptual level reveals a difference of 31% between more concrete and abstract question forms in the picture condition and 34% in the no-picture condition.

Table 4.1
Mean Number of wh- Questions Correct for Each Condition, Conceptual Categories, and the Conditions Combined

	Picture	**No Picture**	**Combined**
Who	87%	92%	89%
What	87%	97%	92%
Where	84%	87%	84%
Who + What + Where	**86%**	**92%**	**89%**
When	33%	53%	43%
Why	63%	51%	57%
How	68%	69%	69%
When + Why + How	**55%**	**58%**	**56%**

Note. Conceptual groups (who + what + where; when + why + how) are bolded.

The descriptive statistics reveal important information regarding the question types across conditions. To provide additional information regarding these differences, a 6 X 2 logistic regression model with multiple observations between subjects was constructed using these responses. This logistic regression was selected to determine if there were differences in the probability of correctly answering questions with different wh- question words and the effect of the picture versus no picture condition on answering these questions. The probabilities of answering questions correctly were modeled on differences between the wh-question words, whether the questions were asked in the picture or no-picture condition, and

the interaction between the two. The coefficients and standard errors for this model are included in Table 4.2.

The coefficients reported in Table 4.2 reveal the parameter estimates that were in the logistic regression model used for analysis. To construct these parameter estimates, first each wh- question word was compared to *why* in the no picture condition. In the calculation of these parameter estimates, *why* was chosen arbitrarily. Each question word in each condition could have been used as the referent to obtain parameter estimates. The comparison revealed statistically significant differences between *why* in the picture condition and *who*, *where*, *what*, and *when* in the picture condition. Additionally, the interaction between each question word and the no-picture condition was compared to the interaction between the no-picture condition and answering *why* correctly. Essentially, this tests the null hypothesis that the picture and no-picture effect is the same for *why* as it is for the other wh- question words. In the case of *what* and *when*, the effect of referential condition is statistically significant. This is not the case for the other question words.

Table 4.2 also displays odds ratios for each of the comparisons between *why* and the other question words in the picture condition. The odds ratios are modeled based on the probability of getting an item incorrect given performance on the *why* picture item. In looking at the question words in the picture condition, the odds ratio of 3.43 for the *when* question reveals that the odds of getting a *when* question wrong in the picture condition are 3.43 times the odds of making an incorrect response to the *why* question in the picture question. For all other question words, an odds ratio less than 1 indicates that there are lower odds of answering those words correctly in the picture condition than *why* in the picture condition. Additionally, the odds ratios of the interactions indicate that for all questions there

are lower odds of the same interaction effect than is seen for the *why* question. The odds are the lowest for *what* and *when* which speaks to the similarity of the interaction effect for those three question words and the referential condition.

Table 4.2
Summary of the Parameter Estimates of the Logistic Regression Model with the Referential Condition, the Difference between wh- Question Words, and the Interaction between Question Words and Referential Condition

	β (SE)	**OR (95% CI)**
(Intercept)	-.54 (.333)	.58 (.30-1.13)
No Picture[a]	.49 (.32)	1.62 (.87-3.04)
How[b]	-.23 (.46)	.79 (.32-1.95)
What[b]	-1.32 (.57)*	.27 (.09-.82)
When[b]	1.23 (.43)*	3.43 (1.49-7.91)
Where[b]	-1.14 (.53)*	.32 (.11-.90)
Who[b]	-1.32 (.51)*	.27 (.10 - .73)
How x No Picture[c]	-.53 (.48)	.59 (.23-1.50)
What x No Picture[c]	-2.1(1.12)*	.11 (.01-.98)
When x No Picture[c]	-1.18 (.53)*	.31 (.11-.88)
Where x No Picture[c]	-.92 (.56)	.40 (.13-1.20)
Who x No Picture[c]	-1.09 (.64)	.34 (.10-1.18)

Note. SE = Standard Errors. OR = Odds Ratio. CI = Confidence Interval. [a] Reference category is Picture. [b] Reference category is Why. [c] Reference category is Why x No Picture. *p < .05.

Parameter estimates were calculated to conduct analysis under the logistic regression model in subsequent analysis. With this model, there was an overall statistical significance between the probability of answering different questions correctly, χ^2 (5) = 61.21, p < .001. There was not, however, an overall statistically significant finding when looking at the probability of correctly answering different wh- questions based on whether the questions were asked in the picture and no picture conditions, χ^2 (1) = 3.54, p = .060. While not statistically significant, this relationship was not insubstantial. Further, there was no interaction between the different wh- questions and conditions, χ^2 (1) = 9.78, p = .082.

A 2 x 2 logistic regression model based on this original model with multiple observations between subjects was used to examine if differences existed when the question words were placed into groups based on their conceptual level (e.g., *who, what, where* and *when, why, how*). The use of these models allowed for an examination of the estimated probabilities of answering a question correctly when: (1) individual wh- question words; (2) groups of wh- question words (e.g., *who, what where* picture vs. *when, why, how* picture); and (3) individual questions words across conditions (e.g., *who* picture vs. *who* no picture) were contrasted with one another. The first set of contrasts examined whether the picture referent had any effects on individual pairs of wh- question words (e.g., *who* picture vs. *who* no picture) or on the conceptual groups of wh- question words (e.g., *who, what, where* picture vs. *who, what, where* no picture). The results of these contrasts are presented in Table 4.3.

Table 4.3
Chi-Squared Values (probability values) of Contrasts between Picture and No-Picture Conditions.

Who	What	Where	When	Why	How	Who, What, Where	When, Why, How
1.44 (.263)	2.95 (.089)	.96 (.332)	4.20 (.044)	2.43 (.126)	.02 (.890)	3.29 (.060)	.15 (.609)

The results of each of the contrasts indicates that similar to the overall analysis, there were generally no statistically significant findings in the responses to questions asked in each of the conditions. The only exception to this was in response to the *when* question. In this case, students answered the question more successfully in the no-picture condition. There were also no statistically significant findings between the picture and no-picture conditions for the conceptual groups of questions. Although not statistically significant, the resulting p-value contrasting the *who, what,* and *where* picture and no picture condition was not

insubstantial, χ^2 (1) = 3.55, p=.060. However, the estimated probability of a correct answer was higher in the no picture condition than the picture condition (.94 vs. .86). Interpretation of these results leads to the rejection of the overall hypothesis that questions would generally be comprehended more successfully in the picture condition. As a whole, the picture did not appear to help students answer the more abstract questions significantly more successfully than when they were presented in the no-picture condition. The only contrast of statistical significance was between the *when* picture and no-picture conditions, and this worked in the opposite direction of what was hypothesized.

The 2 x 2 logistic regression model was also employed to determine if there were differences between answering questions based on their conceptual levels. The conceptual groups of questions were compared to determine if the contrast between the probability of answering one conceptual category of questions and another conceptual category was statistically significant. The conceptual categories of questions were compared within their condition (e.g., *who, what, where* picture vs. *when, why, how* picture) and when combined across conditions. The results of this analysis are presented in Table 4.4.

Table 4. 4
Values of Contrasts between Conceptual Groups

	Who, What, Where	When, Why, How		
	Probability Estimate	Probability Estimate	χ^2	*p*
Combined	.91	.56	46.77	< .001
Picture	.86	.55	25.82	< .001
No Picture	.94	.57	29.71	< .001

These results suggest that there are differences between answering questions based on their conceptual level. This was true within each condition and when all questions were

combined. Additionally, each question word contrast was compared within each condition (e.g., *who* + picture vs. *how* + picture). These results are presented in Tables 4.5 and 4.6.

Table 4.5
Chi-Squared Values [probability values] of Contrasts of Individual wh- Question Words in the Picture Condition

Question Word	What	Where	When	Why	How
Who (.86)	0 [1.00]	.084 [.743]	30.47 [.001]	7.49 [.005]	6.08 [.013]
What (.86)		.084 [.743]	36.27 [.001]	5.99 [.013]	4 [.046]
Where (.84)			40.64 [.001]	5.10 [.023]	4.55 [.036]
When (.33)				9.57 **[.002]**	14.47 [**< .001]**
Why (.37)					.24 [.610]

Note. The probability estimates under the model are in parentheses. Bolded numbers indicate unexpected results

Table 4.6
Chi-Squared Values [probability values] of Contrasts of Individual wh- Question Words in the No-Picture Condition

Question Word	What	Where	When	Why	How
Who (.92)	.92 [.318]	.025 [.625]	25 [< .001]	21.22 [< .001]	8.69 [.004]
What (.97)		1.84 [.169]	27.89 [< .001]	31.47 [< .001]	11.11 [.001]
Where (.89)			24.37 [< .001]	18.65 [< .001]	9.47 [.002]
When (.50)				.083 [.871]	5.37 [**.017]**
Why (.51)					4.38 **[.036]**

Note. The probability estimates under the model are in parentheses. Bolded numbers indicate unexpected results

These contrasts were expected to fall along conceptual lines. For example *who* + picture was expected to be statistically significantly different from *why* + picture, but *who* + picture would not be expected to differ significantly from *what* + picture. Unexpectedly, *when* was different from all other question in the picture condition and *how* in the no picture condition. Additionally, *why* was different than *how* in the no picture condition.

In summary, examination of the descriptive statistics reveals that students answered all questions except for *why* questions better in the no picture condition than the picture condition. There was not a statistically significant finding for whether the questions were presented in the picture or no picture condition. There were, however, some potentially meaningful, if not significant, differences between the picture and no picture condition for some of the wh- question words. Additionally, there was a statistically significant difference between the success the participants had answering concrete and abstract questions. This was consistent across conditions.

The Influence of Receptive Vocabulary Ability on the Comprehension of Wh- Questions

The second research question addressed the relationship between receptive vocabulary and question comprehension across the two conditions. Specifically, the second research question and its sub-questions were:

What is the relationship between receptive vocabulary and question comprehension in school age individuals with ID?

a. Is there a relationship between overall wh- question comprehension and receptive vocabulary?
b. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary ability based on whether the question words are more concrete or abstract?
c. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary based on whether the questions are asked in the picture or no-picture condition?

d. Is there a difference in the relationship of wh- question comprehension and receptive vocabulary ability based on whether the question words are more concrete or abstract within the picture and no-picture condition?

The first step of the analysis relative to this set of questions was to enter PPVT- 4 (Dunn & Dunn, 2007) raw scores into both logistic regression models as a covariate to determine their relationship to the probability of correctly answering wh- questions. The parameter estimates for the model are reported in Table 4.7.Similar to the model without the PPVT-4, the *why* question in the picture condition is used as a referent. Additionally, the interaction between the question word and the no-picture condition was compared against the interaction of *why* and the no-picture condition. This interaction examined whether the effect of the picture or no-picture condition for the *why* question was different than the picture/ no-picture effect for the other wh- questions. The only difference between this model and the first model is that the PPVT-4 is entered as a covariate.

Similar to the previous model, all question words in the picture condition had statistically significant values when compared to *why* in the picture condition. With the PPVT-4 (Dunn & Dunn, 2007) entered into the model, the only interaction contrast that was statistically significant was the comparison of the effect of the referential condition on *when* versus *why*. Also similar to the previous model is that the odds of getting a *when* question incorrect in the picture condition are greater than the odds of getting a *why* question incorrect in the picture condition. The odds ratio under one for all other question words indicates that there are lower odds of getting a question in the picture condition incorrect when compared to *why* in the picture condition. Additionally, entering the PPVT-4 (Dunn & Dunn, 2007)

results in similar odds as the previous model, with there being an interaction effect between referential condition and the question word. Here, all question words had lower odds of having a different interaction effect than *why*. Again, *what* and *when* had the lowest odds of having a different interaction effect than *why*.

Table 4.7
Summary of the Parameter Estimates of the Logistic Regression Model with the Referential Condition, the Difference between wh- Question Words, the Interaction between Question Words and Referential Condition, and the PPVT-4 Raw Scores

	β (SE)	**OR (95% CI)**
(Intercept)	-2.04 (.82)	7.70 (1.53-38.78)
No Picture[a]	.56 (.35)	1.8 (.90-3.57)
How[b]	-.25 (.52)	.78 (.28-2.16)
What[b]	-1.46 (.62)*	.23 (.07-.78)
When[b]	1.37(.48)*	3.95 (1.55-10.06)
Where[b]	-1.27 (.58)*	.32 (.11-.90)
Who[b]	-1.38 (.55)*	.28 (.09 - .88)
PPVT-4 Raw	-.03 (.01)*	.97 (.96-.99)
How x No Picture[c]	-.60 (.52)	.55 (.20-1.52)
What x No Picture[c]	-2.26(1.18)	.11 (.01-1.05)
When x No Picture[c]	-1.37 (.59)*	.25 (.08-.80)
Who x No Picture[c]	-.95 (.62)	.39 (.12-1.31)
Where x No Picture[c]	-1.22 (.64)	.30 (.08-1.05)

Note. SE = Standard Errors. OR = Odds Ratio. CI = Confidence Interval. [a] Reference category is Picture. [b] Reference category is Why. [c] Reference category is Why x No Picture. *p<.05.

In the 6 x 2 logistic regression model, an overall statistically significantly relationship between the PPVT- 4 raw scores and the probability of answering questions correctly was found, χ^2 (1) = 12.96, p < .001. This showed that there was an overall positive relationship between receptive vocabulary ability and the probability of answering questions correctly. Additionally, the logistic regression models were used to examine relationships between

receptive vocabulary, the conditions in which the pictures were asked, and the conceptual level of the questions.

Similar to the logistic regression models without the PPVT- 4 (Dunn & Dunn, 2007) raw scores entered in as a covariate, the difference between answering different types of questions was statistically significant, χ^2 (1) = 12.96, p < .001. The overall difference between answering the questions based on whether they were presented in the picture or no picture condition was not statistically significant, χ^2 (1) = 3.24, p=.072. There was still no interaction between the condition in which the questions were presented and the overall questions answered correctly, χ^2 (1) = 9.53, p=.090. Similar to the model without the PPVT-4 as the covariate, the difference between the picture conditions and the interaction between the wh- question and referential condition were not statistically significant, yet they were not insubstantial.

These logistic regression models were also used to more closely examine differences between the individual question words and conceptual groups of questions based on condition. The results are presented in Table 4. 8.

Table 4.8
Chi-Squared Values (probability values) of Contrasts between Picture and No-Picture Conditions with the PPVT- 4 as a Covariate.

Who	What	Where	When	Why	How	Who, What, Where	When, Why How
1.42 (.210)	2.68 (.111)	.79 (.422)	4.46 (.037)	2.85 (.089)	0 (.975)	3.27 (.070)	.15 (.698)

In terms of statistical significance, the results of these contrasts were the same as the model without the PPVT-4 (Dunn & Dunn, 2007) entered as covariate. The probability of answering *when* correctly was statistically different between picture conditions. In this model

as well, there was a greater probability of answering a *when* question correctly in the no-picture condition than the picture condition. These results suggest that there is not a difference in the comprehension of wh- question words based on whether they are presented in the picture or no-picture condition with the exception of *when*. Investigation of the probability estimates for *why* questions reveals that students were more likely to answer these questions in the picture condition; however, the probability did not differ statistically across conditions. For all other questions, the picture did not appear to support the understanding of questions even when the PPVT- 4 was entered as a covariate. Interestingly, both *why* and *when* are members of a conceptual group (*when*, *why*, *how*) that did not vary significantly when the picture and no picture conditions were compared. Similar to the model without the PPVT-4, the statistical results of the hypothesis test between the concrete questions (*who, what, where*) is not significant, but not trivial.

Additionally, the PPVT- 4 (Dunn & Dunn, 2007) raw scores were entered into these logistic regression models as a covariate to determine if they impacted the statistical significance of the conceptual groups as a whole. They did not. Results are in Table 4.9.

Table 4.9
Values of Contrasts between Conceptual Groups with the PPVT- 4 Entered as a Covariate

	Who, What, Where	When, Why, How		
	Probability Estimate	Probability Estimate	χ^2	*p*
Combined	.92	.56	45.36	< .001
Picture	.88	.55	29.57	< .001
No Picture	.92	.56	28.15	< .001

Individual wh- question contrasts were also examined to determine whether they fell along conceptual lines when the PPVT- 4 (Dunn & Dunn, 2007) raw scores were entered into the model. These results are shown in tables 4.10 and 4.11.

Table 4.10
Chi-Squared Values [probability values] of Contrasts of Individual wh- Question Words in the Picture Condition with the PPVT- 4 Raw Scores as a Covariate

Question Word	What	Where	When	Why	How
Who (.88)	0 [.912]	.02 [.863]	32.65 [<.001]	6.68 [.010]	6.25 [.015]
What (.89)		.09 [.763]	41.02 [<.001]	6 [.015]	4 [.041]
Where (.87)			42.87 [<.001]	5.04 [.026]	5 [.030]
When (.32)				9.69 **[.002]**	14.73 [**< .001]**
Why (.65)					.20 [.636]

Note. The probability estimates under the model are in parentheses. Bolded numbers indicate unexpected results

Table 4.11
Chi-Squared Values [probability values] of Contrasts of Individual wh- Question Words in the no-Picture Condition with the PPVT- 4 Raw Scores as a Covariate

Question Word	What	Where	When	Why	How
Who (.93)	.79 [.327]	.03 [.619]	22.78 [< .001]	20.40 [< .001]	8.27 [.005]
What (.98)		1.56 [.198]	26.10 [< .001]	29.87 [< .001]	10.09 [.001]
Where (.91)			23.80 [< .001]	18.50 [< .001]	8.91 [.003]
When (.51)				0 [.996]	5.17 [**.025]**
Why (.51)					4.83 **[.030]**

Note. The probability estimates under the model are in parentheses. Bolded numbers indicate unexpected results

In the picture condition *when* was the only question word with unexpected results. Here, the probability of answering a *when* question in the picture condition was significantly different from all other question words. In the no-picture condition, *why* was significantly different from *how*. The results of all other contrasts were expected based on the conceptual category of the question. Based on this and the group evidence, there is an overall difference

between conceptual groups of questions when accounting for the receptive vocabulary levels of the participants and this difference is similar across conditions.

Additionally, the differential relationship between receptive vocabulary ability and the ability to answer questions based on their conceptual level was explored within and across conditions. Using the logistic regression models with the PPVT- 4 as a covariate, estimates of probabilities for each of the individual question words and conceptual groups were calculated. A model was created in which the probabilities were estimated based on whether the PPVT- 4 raw score fell within or outside a standard deviation of the mean of the sample. Table 4.12 presents estimates of the probabilities of answering the conceptual groups of questions correctly based on receptive vocabulary ability. The estimates are presented for each condition and when the conditions were combined.

These estimates of probabilities reflect the difference between the groups of questions and how they relate to the receptive vocabulary abilities of the participants. Inspection of the results reveals a clear difference in the relationship between receptive vocabulary ability and the two conceptual levels of the questions. Specifically, the receptive vocabulary appears to have a weaker relationship with *who, what, where* questions than it does with *when, why, how* questions. Although there is an increase in the estimates of probabilities between the PPVT- 4 raw score with the concrete questions (*who, what, where*), this difference is greater with the abstract questions (*when, why, how*). In the picture condition there is a greater difference across the continuum of modeled probabilities (low to high receptive vocabulary) for the concrete questions than in the no-picture condition. Although this is the case, there is a clear difference in the relationship of the PPVT-4 to the two groups of questions across conditions. The graph in Figure 4.1 visually illustrates this difference when the conditions

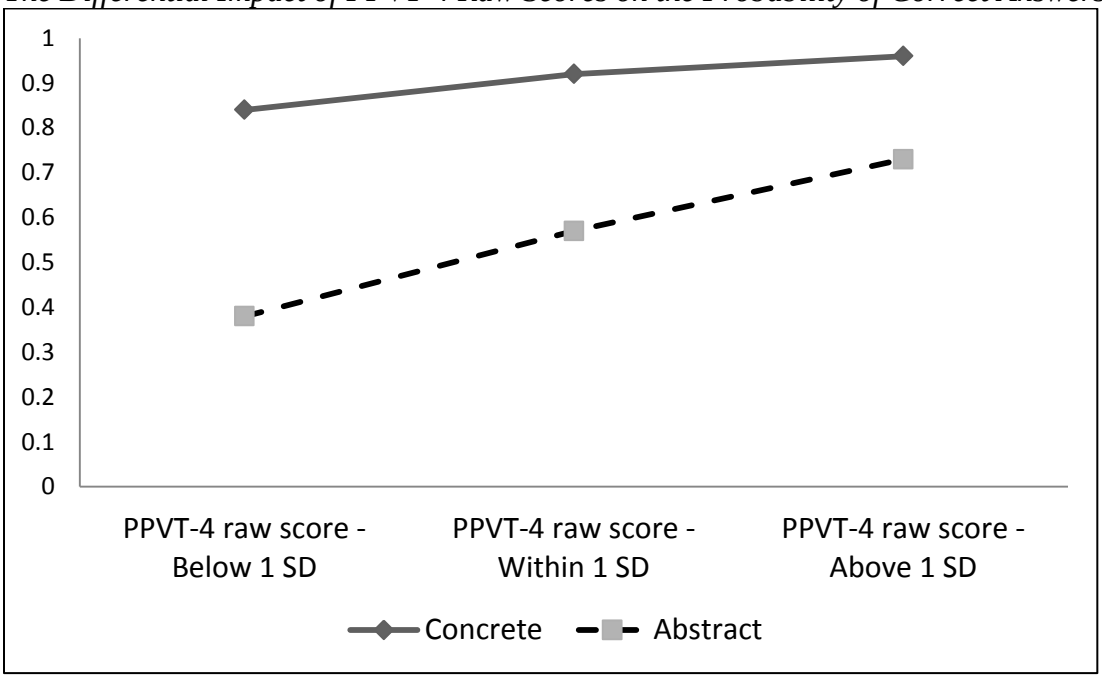
were combined. These results show that receptive vocabulary ability is differentially related to the answering of different conceptual groups of questions.

Table 4.12
Estimates of the Probability of Correct Responses for Groups of Answers with the PPVT- 4 Entered as a Covariate

Groups and Condition	1 SD below	Within 1 SD	1 SD above	Difference
Who, What, Where Combined	.839	.917	.959	.117
When, Why, How Combined	.377	.562	.731	.354
Who, What, Where Picture	.770	.876	.938	.168
When, Why, How Picture	.369	.553	.724	.355
Who, What, Where No Picture	.891	.945	.973	.082
When, Why, How No Picture	.385	.571	.738	.353

Note. SD = Standard Deviation. The numbers in the columns refer to the modeled probabilities of answering a question correctly for participants who would fall below, within, and above 1 standard deviation of the sample mean on the PPVT- 4 raw scores. Bolded/unbolded rows differentiate the conceptual groups.

Figure 4.1
The Differential Impact of PPVT-4 Raw Scores on the Probability of Correct Answers



The results of these analyses indicate that the PPVT-4 raw score was related to the ability of the participants to answer questions. This was the case in each condition. When the PPVT-4 was entered into the logistic regression model as a covariate, there was still no statistically significant finding between the picture and no picture conditions. With the PPVT-4 entered as a covariate, there was still a statistically significant finding when the probability of answering concrete and abstract questions was compared. This was the case regardless of condition. The PPVT-4 raw score was also shown to have a greater impact on the probability of answering abstract questions as opposed to concrete questions across both conditions.

Error Analysis

The third research question and its sub-questions investigated the type of errors participants made when responding to the different question types. Specifically, research question 3 reads,

> What types of responses do school age individuals with ID give when their response to a question is incorrect?
>
> a. Do patterns of errors exist in the incorrect responses school-aged individuals with ID provide (i.e.., categorical relatedness, substitutions, I don't know/no responses, topic related, unrelated)?
> b. Do errors or patterns of errors in the incorrect responses of school-aged individual with ID vary based on the use of a picture referent?

The first step in answering these questions involved categorizing incorrect responses based on the researcher-developed coding scheme (see Appendix B). Through these procedures incorrect responses were categorized as follows: categorically related, substitutions, I don't know/no response, and repetitions. The remaining errors were divided

into categories indicating that they were either related to lunch or the picture or unrelated. Incorrect responses were then grouped in relation to the condition the question was asked in, conceptual category, and the characteristics of the response. The results of the answers using this scheme are presented in Tables 4.13 and 4.14 relative to the conceptual category of the questions.

Table 4.13
Frequency of Errors by Type for *Who*, *What*, and *Where* Questions.

	Total Errors	Categorically Related	Sub	IDK/NR	Rep	Topic Related	Unrelated
Who - P	5	1 (20%)	1 (20%)	0 (0%)	0 (0%)	3 (60%)	0 (0%)
Who - NP	3	0 (0%)	0 (0%)	0 (0%)	1 (33%)	1 (33%)	1 (33%)
Who Combined	**8**	**1 (13%)**	**1 (13%)**	**0 (0%)**	**1 (13%)**	**4 (50%)**	**1 (13%)**
What - P	5	2 (40%)	0 (0%)	1 (20%)	1 (20%)	1 (20%)	0 (0%)
What - NP	1	0 (0%)	0 (0%)	1(100%)	0 (0%)	0 (0%)	0 (0%)
What Combined	**6**	**2 (33%)**	**0 (0%)**	**2 (33%)**	**1 (17%)**	**1 (17%)**	**0 (0%)**
Where- P	6	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (83%)	1 (17%)
Where - NP	5	0 (0%)	1 (20%)	1(20%)	0 (0%)	2(40%)	1(20%)
Where Combined	**11**	**0 (0%)**	**1 (9%)**	**1 (9%)**	**0 (0%)**	**7 (64%)**	**2 (18%)**

Note. P = Picture; NP = No Picture; IDK/NR = I don't know/ No response; Sub = Substitution; Rep = Repetition. Percentages of response type of total errors per category are in parentheses.

There were few incorrect responses to *who*, *what*, and *where* questions. This makes it difficult to make generalizations about the responses and therefore the results should be interpreted with caution. However, there are notable differences in the number of errors in the *what* picture condition as opposed to the no picture condition, and a large portion of errors in response to the *where* questions were topic related.

Table 4. 14
Frequency of Errors by Type for *When*, *Why*, and *How* Questions.

	Total Errors	Categorically Related	Sub	IDK/NR	Rep	Topic Related	Unrelated
When - P	26	9 (33%)	6 (21%)	5 (19%)	2 (8%)	3 (12%)	1 (4%)
When - NP	18	8 (47%)	2 (12%)	2 (12%)	1 (6%)	3 (24%)	2 (6%)
When Combined	**44**	**17 (39%)**	**8 (18%)**	**7 (16%)**	**3 (7%)**	**6 (14%)**	**3 (7%)**
Why - P	14	4 (29%)	0 (0%)	1 7%)	1 (7%)	4 (29%)	4 (29%)
Why - NP	18	3 (17%)	3 (17%)	3(17%)	2 (11%)	3 (17%)	4(22%)
Why Combined	**32**	**7 (22%)**	**3 (9%)**	**4 (13%)**	**3 (9%)**	**7 (16%)**	**8 (25%)**
How - P	12	2 (17%)	0 (0%)	1 (8%)	0 (0%)	6 (50%)	3 (25%)
How - NP	11	2 (18%)	0 (0%)	3(27%)	2(18%)	2(18%)	2(18%)
How Combined	**23**	**4 (16%)**	**0 (0%)**	**4 (17%)**	**2 (9%)**	**8 (35%)**	**3 (13%)**

Note. P = Picture; NP = No Picture; IDK/NR = I don't know/ No response; Sub = Substitution; Rep = Repetition. Percentages of response type of total errors per category are in parentheses.

In contrast, there were substantially more errors in response to the *when*, *why*, and *how* questions. *When* questions were the most difficult for the students in the study to answer; however, these errors were more likely to be categorically related than in response to any other question type. Interestingly, there was a relatively high prevalence of topic-related answers in the *how* picture condition. Overall, participants had similar numbers of categorically related incorrect responses in the picture condition as in the no-picture condition indicating that the picture did not scaffold the ability to provide categorically-related but incorrect answers. There were substantial differences in terms of categorical relatedness in the percentage of picture + *who, what* and *where* questions versus no picture + *who, what*, and *where*. This difference may have been due to the fact that there were more acceptable answers for the no-picture than the picture condition (e.g., any food item would

have been correct in the no picture condition, but students needed to provide food items in the picture item in the picture condition). The same differences did not exist between the *when*, *why*, and *how* questions.

These numbers suggest that when students do not know the answer to the questions they generally respond in a way that matches one of the classifications. Only 15% of the answers did not fit into one of the other categories. This percentage was higher in the no-picture condition than in the picture condition (18% vs. 10%), indicating that students may have used the picture as a type of scaffold when they did not understand the question. Further, it is possible that this percentage was higher than it might have been had the inclusion criteria for the substitution category not been so strict (i.e., if categorically related responses were included as substitutions). These results do indeed suggest that there are identifiable error patterns that exist when students with intellectual disabilities do not understand wh- questions.

Summary of Findings

The results of the analysis conducted in the current study reveal several important findings regarding wh- question comprehension in individuals with ID. The first important finding is that individuals with ID comprehend different wh- questions with varying levels of success. Furthermore, success answering these questions appears to be related to whether the questions are more concrete or more abstract in nature. The overall probability of correctly answering questions with different wh- question words is related to receptive vocabulary, which appears to have more of an impact on the probability of answering more abstract questions correctly than correctly answering more concrete questions.

No substantial overall differences were found in answering questions that were asked in a picture versus a no-picture condition with and without receptive vocabulary ability used

as a covariate. When receptive vocabulary ability was accounted for, there was a statistically significant finding when comparing the probability of answering the *when* question in both conditions. Further, although the analyses did not reveal significant differences among other question words, some of these results approached significance. For example, in both logistic regression models, the differences between answering *what* and *why* questions in each condition deserve further exploration. In the case of *what*, students answered more questions correctly without the picture while for *why*, students answered more questions correctly about the picture.

Error analysis of the participants' incorrect answers also indicated that there were distinct, definable categories that could account for approximately 85% of their incorrect responses regardless of the condition in which the questions were asked. Furthermore, 25% of responses were answered with a response that was categorically related to the question word and 10% were answered by substituting a wh- question word they already comprehended. Participants did produce a greater percentage of incorrect answers related to the topic in the picture condition than in the no-picture condition. It was difficult to interpret differences in error patterns between the answers to conceptual groups of questions because there were substantially fewer incorrect responses to concrete questions than more abstract questions.

CHAPTER 5

Discussion

This study investigated how different wh- question words and referents influenced the comprehension and answers of wh- questions in individuals with intellectual disabilities (ID). The findings of logistic regression analysis indicate that the probability of answering a wh-question correctly varied based on the wh- question word that was used. These differences appear to be based along conceptual lines as participants were more successful answering *who*, *what*, and *where* questions as opposed to *when*, *why*, and *how* questions. Further, receptive vocabulary ability appeared to be linked to the success of answering wh- questions correctly. The relationship between receptive vocabulary and the probability of correctly answering questions differed based on the conceptual level of the questions. The use of a picture referent did not make an overall difference in the probability of answering questions correctly, but did impact the ability to answer some wh- questions. Analysis of students' incorrect answers revealed that a majority of responses could be placed into distinct categories. These findings will be interpreted and discussed in relation to previous research and the research questions that guided the current investigation.

Relative Ease in Answering Wh- Questions

There were differences in the ease of answering questions with different wh- question words across both conditions. The rank ordering of correct answers when all conditions were combined was: *what*, *who*, *where*, *how*, *why*, *when*. This rank order was the same in the picture condition, and in the no-picture condition, the first four were in the same rank order

but the last two, *when* and *why*, were reversed. When the differences between the probability of answering different wh- questions correctly were investigated more closely, the probability of answering questions that were more concrete (began with *who*, *what*, and *where*) was higher than those that were more abstract (began with *when*, *why*, and *how*). This was the case in both referential conditions. As is true with typical development and more distinctly in individuals with language impairment (e.g., Ervin-Tripp, 1970; Lee & Ashmore, 1983; Tyack & Ingram, 1977), these participants with ID found questions with more concrete question words easier to understand than those with more abstract words.

Influence of Conceptual Level on the Relative Ease of Answering Questions

Comprehension of *when*, *why*, and *how* questions often occurs later in typical development probably because questions that begin with these question words are conceptually more complex. *How*, *when*, and *why* questions require understanding of the concepts of manner/means, temporality, and causality (Tyack & Ingram, 1977). This may be why students with language impairments and intact cognitive ability find these question types difficult to understand (Lee & Ashmore, 1983). It may also explain the contrast between the probability of answering more concrete versus abstract questions in the current study. These findings regarding different types of question words are also similar to the difficulty some children with autism and ID experience responding to inferential versus concrete questions (Hewitt, 1998). In addition to varying in their conceptual level, the two groups of questions differed in their level of abstraction. For example, the answers to *who*, *what*, and *where* questions in the picture condition were explicitly depicted in the picture. Even when a referent was not immediately present in the no-picture condition, the answers were concrete nouns in the case of *who* and *what* and often a prepositional phrase or noun in the case of the *where* question. These are all things that could be visualized and are very familiar to the

participants. For *when*, *why*, and *how* questions, the answers required students to go beyond physical referents and apply more abstract thinking. In the picture condition, these three questions required students to make inferences. Only the *how* question had a salient cue provided in the picture (the lunch tray) and even it required an inference (i.e., if the boy is eating his lunch from a lunch tray he must have purchased or received free/reduced lunch from the cafeteria). These differences may have been partly responsible for the difference between the two groups of questions, and they provide important additional information regarding the factors that may impact the success students with ID have in answering different types of questions.

The Relationship of Syntax to the Relative Ease of Questions

Given the delays that are typically seen across receptive and expressive syntax skills in individuals who have ID (e.g., Paul, 2007), it is unsurprising that the participants in the study had more difficulty with *when*, *why*, and *how* questions. It was speculated that the expressive language difficulties often seen in individuals with ID would lead to difficulty formulating responses for questions such as *how* and *why*. This held true for *why* questions because correct responses involved combining different parts of speech (e.g., "so he won't starve" and "him hungry"). However, acceptable answers to *how* questions included single word responses consisting solely of a noun (e.g., "money," "cafeteria," "lunch-lady"). Some students replied with more complete answers such as "He go to the cafeteria." Indeed, across conditions students were more successful answering *how* questions (68%) than *why* questions (59%), although the percentages were similar in the picture condition (68% for *how* vs. 66% for *why*). It is possible that the differences were due to difficulties with expressive language. *How* questions were shown to be the most difficult for students with language disorders (Lee & Ashmore, 1983), and in some studies with participants who were typically developing

(Tyack & Ingram, 1977; Cairns & Hsu, 1978). These studies may have required more complete answers in order to be considered correct. For example, Cairns and Hsu (1978) required grammatically appropriate responses in order to receive complete credit for a correct response.

Another interesting point about the *how* question is that many of the responses would have been correct responses if the *how* in the question was substituted with a *where* (i.e., "*Where* does the boy get lunch?" rather than "*How* does the boy get lunch?"). Examples of these correct answers include "from the cafeteria," "you got to go to the cafeteria," and "at cafeteria." It is possible that students used a substitution strategy for these types of questions. In all, 10 of the correct responses in the picture condition and 9 in the no-picture condition could have been acceptable if *where* was substituted for *how*. All of these students answered the *where* question in each relative condition correctly. Although the students had a greater probability of success answering these questions than *when* in both conditions and *why* in the no picture condition, the fact that *how* was still significantly different than all of the concrete question words provides more evidence that correctly answering these two different types of questions may require different skills (e.g., inferencing) that may not be directly related to expressive syntax ability.

Additionally, the combination of syntactic and semantic qualities of *who*, *what*, and *where* questions versus *when*, *why*, and *how* questions may have contributed to the differences between the two types of questions. *Who*, *what*, and *where* questions function like pronouns for the sentence constituent they replace (e.g., *what* refers to cookie in "The boy eats a cookie"). In contrast, the *when*, *why*, and *how* questions are more complex in that the reason, manner, or time refers to information that is encoded either through semantic

relations within a sentence or an entire clause (Bloom et al., 1982; Owens, 2008). As the ability to use these types of questions typically occurs later in development, it is reasonable to assume that understanding and subsequent expression of answers might be somewhat related to the same factors. The language difficulties of many individuals with ID may have contributed to the differences seen between these types of questions.

Differences Between the Picture and No-Picture Condition

Across the picture and no-picture condition, there were no statistically significant differences in the probability of answering questions correctly. There was, however, a statistically significant finding in the probability of answering a *when* question correctly based on condition. The overall finding was surprising as previous research by Parnell and colleagues (1983, 1984) highlighted developmental differences based on whether a referent was immediately present or not for children who were typically developing. Further, statistically significant differences were found between children who were typically developing and those with language delays based upon whether a referent was present or not (Parnell et al., 1986). Although these studies led to the belief that differences would be seen based on condition, variations in methodology may have led to the dampening of the potential effect of the referential condition.

Overall Effects of the Referential Conditions on Question Comprehension

One of the key differences between the current study and previous studies that have used pictures or storybooks (e.g., Ervin-Tripp, 1970; Tyack & Ingram, 1977) as stimuli for question answering was that one picture was used as the referent for all of the questions in the current study. Because the participants in the current study had ID and language delays, one picture was selected to reduce the possibility of introducing uncontrolled variability with each picture. Using multiple pictures would have potentially introduced different characters, vocabulary, and contextual situations.

In that case, participant knowledge of the vocabulary and context within the picture would have been accounted for and measured. By using a single picture about a familiar topic this variability was avoided.

Factors related to the use of this single picture may have led to the fact that there were no overall statistically significant differences between the picture and no-picture condition. It was hypothesized that picture support would have helped students answer questions more successfully than no picture questions because additional context was provided. However, the selection of an everyday routine in the no-picture condition, a topic students had a great deal of knowledge about and may have had practice answering (e.g., parents may have asked these questions), may have eliminated this potentially positive effect. It has been speculated that children are better at answering questions about routines than novel situations (Lokusa, Ryder, & Leinonen, 2007). This may have inflated results in the no-picture condition.

Although knowledge of context may have washed out differences between the picture and no-picture referential conditions as a whole, students still may have used the picture to aid in the comprehension of certain abstract questions. In the picture condition, students had to answer questions about "the boy." This may have introduced an added layer of difficulty to the questions that was not present in the no-picture condition. They were required to make inferences for the *when*, *why*, and *how* questions. In the no-picture condition, little inferencing was required as students answered questions regarding their own familiar routine. It was hypothesized that the picture would act as a scaffold to provide the context necessary to help the participants integrate their world knowledge and be more successful answering these questions (Leinonen, Letts, & Smith, 2000). It is possible that students in the study did this for some of the questions as the picture questions may have been more difficult than the no-picture questions. However, it is also possible that students' contextual knowledge

eliminated the need to integrate this knowledge. Both of these may be reasons why as a whole the findings were relatively similar between the referential conditions.

Differences of Individual Wh- Questions in Different Referential Conditions

The disparate results of the when and why questions between referential conditions warrant further investigation. In the picture condition, both of these questions required participants to make inferences. When questions in the picture condition were answered correctly only 36% of the time, while they were answered correctly 53% of the time in the no-picture condition. There was no obvious time cue in the picture such as a clock (and it is unclear whether participants could tell time in order to use the clock if it did exist). Participants were required to make an inference about the students in the picture to answer the question correctly. Three of the participants who answered the question correctly in the no-picture condition (e.g., 12:00) replied with an "I don't know" response when asked the parallel question in the picture condition. Two others replied with "anytime" although they answered the question correctly in the no-picture condition. A reference to a specific time or time of the day was required to be credited with a correct answer in this case. It is possible that these students were confused when trying to answer this question because there was no visual cue, or it may have been that they don't have enough temporal knowledge to understand that lunchtime is a rather universal time for all students. Interestingly, Parnell and colleagues (1984, 1986) found that children who were typically developing and those that had a language impairment answered when questions better when no referent was present than when they were asked about an object. The example of the picture they used in their study for the when question was not about a different character (e.g., "the boy") but was about a routine. The question Parnell and colleagues provided about the object (i.e., "When

did I drop my pencil?") required an immediate temporal response, which children may have more difficulty responding to (Ervin-Tripp, 1970). If the when picture question used in the current study referred to the participant rather than another character (e.g., if it were about the participant's routine), it is possible that they would have answered it correctly a greater percentage of the time.

Alternately, students answered more *why* questions in the picture condition than in the no-picture condition (63% vs. 51%). Here, similar to *when*, there was no visual cue to support the response to this question in the picture referent. However, for the students who answered the question correctly in the picture condition but answered it incorrectly in the no-picture condition, the picture did indeed seem to function as a scaffold. All of these students responded to the question in both conditions, but were able to make the inferential leap with the added context in the no-picture condition than they were in the no-picture condition. Additionally, four of the participants who answered the *why* question incorrectly in the picture condition described something that was in the picture without using a causal word. Further, two of the participants who had a causal word in their incorrect response finished the statement by describing something that was in the picture (e.g., "Because he eat with the girl"). This indicates that the picture had a different effect on the responses for the *why* question than the other questions.

In terms of the concrete questions (*who*, *what*, *where*), error analysis reveals that the picture may have also limited some of the answers participants gave. Answers to the all three of these questions were in the picture, but all questions in the no-picture condition were answered with a greater percentage of success than in the picture condition. It may be that the picture made it more difficult to respond because *who*, *what*, and *where* questions in the

picture condition required a specific response related to the picture. In the no-picture condition, however, students had more freedom with their responses. For example, "any kind of food" would have been an acceptable response to the *what* question in the no picture condition, but not in the picture condition. It is very likely that this led to some of the discrepancies between the two conditions.

Another interesting aspect of the differences between responses to questions in the picture and no-picture conditions was the way students used the picture when their responses were incorrect. Some students answered *where*, *why*, and *how* questions by describing what was in the picture. For example, incorrect answers to the *where* question included "he eats the food" and "eating sandwich." These types of responses constituted 28% of the incorrect answers for the *where*, *why*, and *how* picture questions. It is possible that this was a strategy used by students to answer questions when they did not know an answer. This could imply that students with ID know they are supposed to use visual cues like pictures to answer questions correctly, but may need to be taught how to do so for each type of question.

Intervention studies with students who have autism and concomitant ID have generally employed pictures as stimuli for aiding in the comprehension of questions. These studies provide interesting information related to the way students with ID may use pictures to comprehend questions (e.g., Hundert & Delft, 2009; Jahr, 2001; Krantz et al., 1981; Secan, Egil, & Tilley, 1989). The participants in these studies were generally able to learn how to use pictures to answer questions (e.g., Hundert & Delft, 2009; Jahr, 2001; Secan, Egil, & Tilley, 1989). In fact, Hundert and Delft (2009) found that students needed to be trained to answer inferential *why* questions in each of their question conditions (i.e., using picture sequence cards, a verbal story, or general information questions). This may indicate that

students with ID need to be taught how to use pictures to answer questions appropriately. The current study assumed that students knew how to use pictures to answer questions and that a picture would therefore act as a scaffold. Perhaps an alternate reason there were no differences between the conditions was that students need to be taught how to use the referent appropriately.

Relationship of Receptive Vocabulary to Answering Wh- Question Words Correctly

For the participants in the current investigation, the probability of answering wh-questions correctly increased as students' raw scores on the PPVT-4 (Dunn & Dunn, 2007) increased for both concrete and abstract question forms. This relationship between receptive vocabulary and wh- question comprehension is unsurprising. The PPVT-4 is a measure of single-word vocabulary. It is logical that there would be a connection between overall vocabulary and this subset of vocabulary targeted in this investigation: wh- words.

Receptive one-word vocabulary has traditionally been regarded as a strength in individuals with ID across age groups in comparison to language domains such as such as receptive syntax (Chapman, 2006; Facon et al., 2002; Laws & Bishop, 2003). In the current study, receptive vocabulary was found to have a statistically significant, positive relationship with both concrete and abstract groups of questions. Furthermore, the estimated probability of answering concrete questions correctly was similar under a logistic regression model for students who would have low, average, and high PPVT-4 scores (Dunn & Dunn, 2007). In contrast, the estimated probabilities differed across these three groups for more abstract questions. This was true across referential conditions. This indicates that a relative strength in receptive vocabulary ability may be most important when answering the more abstract questions. In different populations of individuals with ID, it is speculated that differences between receptive vocabulary and other domains of language may not exist if measures of

conceptual vocabulary are used as opposed to measures of concrete vocabulary like the PPVT-4 (e.g., Chapman, 2006). If a measure assessing conceptual vocabulary was used in the current study, these differences between question types may have been even more pronounced. That is, students who have lower conceptual vocabulary may have had a lower probability of answering the abstract questions correctly. Those at the higher extreme in terms of conceptual vocabulary may have also answered more abstract questions correctly. Additionally, it is possible that all students would have been able to answer the more concrete questions regardless of conceptual vocabulary ability. Different measures of receptive vocabulary may relate to the probability of answering questions correctly in different ways.

Although differences were not found between the probability of answering questions correctly within both groups of questions, when the PPVT-4 (Dunn & Dunn, 2007) was entered as a covariate into the model, there was still a statistically significant difference between the probability of answering concrete versus abstract questions. This was the same finding in the logistic regression model without the PPVT-4 being entered into the model. In fact, although the PPVT-4 raw scores were statistically significant when entered into the model as a covariate, probability estimates changed very little for each individual question word. Additionally, little changed in terms of the individual contrasts between question words across referential conditions (e.g., *who* picture vs. *when* picture). This is surprising given the variability of PPVT-4 raw scores seen in the sample. It does, however, speak to the real differences between the probability of correctly answering the different types of questions regardless of receptive vocabulary.

Error Analysis in Relation to Previous Studies

An important component of research on question comprehension has been analyzing the types of errors children make when they do not understand a question. Early research concentrated on the analysis of errors in an attempt to predict how students would respond to a question they did not understand (Ervin-Tripp, 1970; Tyack & Ingram, 1977). This early research focused on the transitivity and placement of verbs in relation to comprehension. The current study used an approach more consistent with the next phase of research (Cairns & Hsu, 1978; Parnell et al., 1984) and investigated more detailed aspects of the content of the answers. Coding schemes were designed for the purpose of separating answers that were actually correct from those that were incorrect but indicated understanding of the question type. Additionally, incorrect responses were coded to categorize them as wh- question word substitution errors, topic related, non-responses, or completely unrelated to the question at hand (e.g., Lee & Ashmore, 1983; Parnell et al., 1986). In the current study, the goal was to explore what students with ID did when they did not understand wh- questions.

There were far fewer errors in the concrete questions category than in the abstract questions (25 vs. 99). This made the identification of patterns in the abstract questions more apparent. These results will be analyzed and discussed in relation to previous research in order to examine different strategies students with ID use to answer questions when they have do not understand them.

Categorically Related Responses

In the current study, incorrect answers coded as categorically related were closest to what Parnell et al. (1984) referred to as "functionally accurate" and a combination of what Cairns and Hsu (1978) termed type two and three answers. In the current study, there were key differences in the analysis of the concrete questions as opposed to the abstract questions

due to the differences in the sheer number of errors made in response to each question type. For example, two of the four total categorically related responses in the concrete questions were responses to the what question in which students named foods not present in the picture. In the abstract questions, these responses made up 28% of the responses, with a similar percentage of the responses being found across conditions.

Consistent with previous research involving individuals with language differences (Lee & Ashmore, 1983; Parnell et al., 1986), the participants in this study had the most difficulty answering *when* questions. When compared to errors in response to other wh-questions, *when* questions had the highest percentage of responses that were categorically related. If an answer was unknown, the student provided a categorically related response 39% of the time, which means they included some temporal element in their response. There has also been a high percentage of similar types of responses to *when* questions in other studies. For example, Cairns and Hsu (1978) found that 25% percent of all of the answers to *when* questions in their study used a fixed form such as "now" without giving a specific enough answer to be considered fully correct. In the current study, 22% of all *when* answers could be placed into this category. Further, Lee and Ashmore (1984) found that 22% of the errors in their sample of students with language impairments responded to *when* questions with "patterned" answers (e.g., giving a temporal answer not specifically related to the question). In contrast, Parnell, Amerman, and Harding (1986) found that children with language impairment struggled to produce "functionally appropriate" error responses to *when* questions when compared with children who were typically developing. Although the percentages of these types of responses are difficult to compare across studies due to differences in methodology, the prevalence of these types of responses in typical

development, language impairment, and ID suggests that this may be a common strategy used to answer *when* questions when an answer is unknown.

A relatively high percentage of the incorrect responses to *why* questions were also categorically related (22%). Interestingly, five students in the picture condition used causal words (e.g., because, so), and four of them then added a description of something that was in the picture but still provided an incorrect response. In the no-picture condition, four students used a causal word and three of these added descriptions that had something to do with lunch but were not accurate responses. These types of error responses are prevalent among individuals with language impairments as well. For example, Parnell, Amerman, and Harding (1986) found that approximately a quarter of all responses to *why* questions provided similar categorically related responses. Lee and Ashmore (1984) found that over half of the incorrect *why* responses in their study were also like this. These data and the results from the current study suggest that this may be a strategy individuals with language difficulties use when they do not understand a causal question or how to answer a question appropriately.

When participants in the current study had incorrect responses to *how* questions, they provided fewer categorically related responses than the other two abstract questions. Only two participants provided responses that could be considered categorically related. It was more difficult to define the categorically related responses for this question word than any of the other question words because the manner and means in which someone does something does not necessarily have a type of word that signals it is a member of that class (e.g., "because" for a causal response). In the case of this *how* question, categorically appropriate responses could have included a person, place, or the way in which something was done. Similarly, none of the participants in Lee and Ashmore's (1984) study responded with

patterned *how* responses. Cairns and Hsu (1978) also found that almost all error responses to *how* questions were unrelated. They point to the variety of responses (e.g., locative responses, verb + ing) they received to their *how* questions. It may be the same factors that make it difficult to define what constitutes a categorically related response for a *how* question makes it difficult for individuals with ID to develop a consistent strategy to approach them even when they don't know the answer.

There are education/therapeutic implications for the percentage of categorically related responses for *when* and *why*. As these two question words were the most difficult for students to answer correctly, it is possible that teachers/clinicians may be able to use existing skills related to knowledge of those types of questions to scaffold correct answers. If these responses are to be viewed on a continuum from somewhat correct to more correct, it may also be developmentally appropriate to teach students about category before focusing on correctness (Parnell et al., 1986). Additionally, the use of categorically appropriate and correct answers appears to be developmental (Parnell et al., 1984). The existence of both of these types of answers may be related to other aspects of language in which students with ID show delays.

Substitutions

Substitutions were defined as responses that would have been correct had a different question word been substituted for the one in the target question. In the current study, there were differences in the substitutions students made between concrete and abstract questions. Only two substitutions occurred in the concrete questions. It was expected that students would substitute "easier" wh- question forms they have knowledge of when faced with a wh-question they did not understand (e.g., Ervin-Tripp, 1970, Tyack & Ingram, 1978). It

certainly seems unlikely that they would answer with more difficult forms. Additionally, the fact that there were fewer incorrect answers in the concrete responses would also lead to fewer substitution errors than with abstract questions. Further, the who, what, and where questions varied in terms of their syntactic frame, making substitutions as defined in the current study less likely. For example, no other question word could be substituted for what in the frame "_____ food do you like?"

Substitutions were far more prevalent in the *when* and *why* questions. *Where*, *when* and *why* all used the same syntactic and semantic frame ("_____ do you eat lunch?"). Interestingly, students provided *where* responses to *when* questions on six occasions in the picture condition and twice in the no picture condition. This constituted nearly 18% of all of the errors in the *when* picture question. All but one of the students who gave this type of response to the *when* question answered their conditionally respective *where* questions correctly. Parnell, Amerman, and Harding (1986) found that individuals with language disorders produced substitutions of *when* questions in the form of *what* and *where* responses. Alternately, Cairns and Hsu (1978) and Ervin-Tripp (1970) reported that their typically developing participants often answered *when* questions with *why* responses. None of the questions used in the other studies shared exact syntactic and semantic frames, however. Therefore, it is unlikely that they used the same substitution criteria as the current study. It is telling, though, that the students in the current study used an easier form they showed evidence of understanding when they did not know the answer to the more difficult *when* question.

For the *why* question in the picture condition, there were no substitutions. There was, however, one locative response and two causal responses that would have been categorically

related to *where* and *when* questions, respectively, but they were not correct responses to these two substituted questions. These may have been evidence of some sort of substitution process. In the no-picture condition there were three *where* substitutions. As with the *when* questions, the use of *where* responses for incorrect answers is not surprising given that it is the answer to an easier concrete question. In the *how* questions there were no substitutions. As explained previously, however, there were several correct responses that would have also been appropriate answers if *where* was substituted for *how* in the question. It is possible that some students may have been using a substitution process here although their response was, in fact, correct for *how*.

I Don't Know/No Response/ Repetitions

Another interesting response pattern was the use of "I don't know" responses. These were used in all question types except for who. Most prominently 15% of the responses to the when picture condition were "I don't know." As described before, this may have been used as a strategy for students who did not know how to respond to the when picture question. It is possible that they may have been able to answer the question with additional prompting. As a whole, 13% of the errors to all of the questions were "I don't know" responses. It should be noted that eight students gave "I don't know" responses and of these, three used this response multiple times. Two of these students used this in all of their incorrect responses and one used it primarily although he responded "no" and "I don't" to two other questions. This suggests that at least some students with ID use an "I don't know" response rather than attempting to provide an answer when they think they might be wrong.

Repetitions of part of the question were used as response strategies infrequently (8% of errors); however, they were spread relatively proportionally throughout the error

responses. These were always repetitions of single words in the question and not the question as a whole. Seven of the ten repetitions were produced by two students. Both of these students had autism listed as their primary exceptionality on their Individualized Education Plan. Therefore, the majority of these repetitions may have been a form of echolalia and the students may have been using the repetition to seek additional information or clarification regarding the question. Unfortunately, the protocol did not include a process for repeating, rephrasing, or otherwise supporting students in understanding the question if they sought support or indicated they did not understand.

Topic Relatedness

Analyzing the incorrect responses of the participants in the study revealed a relatively large number of responses that were produced that were related to lunch in some way. These types of responses were given for all question words and conditions with the exception of what no-picture, where there was a single incorrect response. How questions had the greatest proportion of these topic-related responses (35%) although how questions had the fewest errors of any of the abstract questions. In the picture condition, if students did not know the answer to questions, they appeared to use the picture as a scaffold for providing some sort of response. As noted previously, many students described or named something in the picture. In the no picture condition, students also often produced responses either describing an aspect of lunch or naming foods or drinks. These types of responses could have been plausible answers to what questions such as “What is the boy doing?” or “What is the boy eating?” They were not classified as strict substitutions because there was no question with an alternate question word that employed one of those syntactic frames. However, it is possible that these responses could represent examples of instances where students respond

to questions they do not understand by providing answers to questions that would use question words they do understand. This has been identified as a common strategy found in developmental studies of wh- question comprehension (Ervin-Tripp, 1970; Tyack & Ingram, 1977). As the majority of students answered what questions correctly in this study, it would not be surprising that this may have been a strategy for answering questions with unknown question words.

All of the remaining responses that did not fit into this category were placed into the unrelated category. These unrelated responses only accounted for 16% of the total errors. As noted before, some of these responses would have been appropriate categorically related responses for the questions that resulted when students substituted the question word in the target question. This may indicate that their responses may not have been completely "unrelated," but reflect their use of a combination of strategies (substitution + categorical knowledge) that was not accounted for in the current study. This possibility is supported by the fact the majority of students who produced these responses produced more than one of these types of responses.

The fact that so many students were able to provide responses related to the topic may also represent an important strategy that could be employed to help students answer questions that require inferences. This may indicate that some students with ID understand that they have to integrate their prior knowledge and context in some fashion, but need additional scaffolding and instruction in order to make that next step and answer the question correctly.

Limitations

This study is the first systematic examination of the comprehension of questions using different wh- question words in individuals with ID. As such, previous developmental

research exploring wh- question comprehension was used as a model and applied to this population. Each of these studies used different stimuli and probed different wh- question forms. In the current study, these means were evaluated and altered to best suit the target population. In applying what was learned from previous studies to this population, several limitations were encountered.

The first limitation is related to the use of the picture and no picture stimuli. This picture was carefully selected in order to account for students' prior knowledge. However, the fact that students had so much prior knowledge about the topic may have diminished real differences between the referential conditions. In the future, an alternate method of controlling for prior knowledge without choosing such a well-known topic may allow for a more complete investigation of the impact of the different referential conditions. The picture that was selected also influenced student performance because there was no salient visual cue for the *when* picture question, and it limited the *how* question that could be asked. Use of different and/or multiple stimuli with different cues may have provided greater opportunity to examine how these factors may have influenced wh- question comprehension.

The use of a single picture also limited the types and number of questions that could be asked. For example, there are multiple types of *what* and *how* questions that could have been asked. Other researchers have asked different varieties of questions such as "*what* + happened," "*what* + do," and "*what* + be" (Parnell et al., 1984; Parnell et al., 1986). The questions used in the current study were easily split into conceptual categories, but the use of these different types of questions could have revealed additional information about the way individuals with ID comprehend the full range of questions they may hear at home and at

school. The limited number of questions in the current study also influences the generalizability of the findings.

Related to this, all of the question forms sampled in this study were object questions. As individuals with ID have shown mixed results in terms of whether they comprehend subject or object questions better (Joffe & Varloska, 2007) and individuals with specific language impairment have shown preference for certain types of object questions (e.g., Deevy & Leonard, 2004; Friedmann & Novogrodsky, 2011), examining difference in the comprehension of *who* and *what* subject versus object may be useful in the future.

Another limitation of this study was that receptive vocabulary as measured by the PPVT-4 (Dunn & Dunn, 2007) was the only measure of language. As it is known that successful understanding and answering of questions is reliant on semantic and syntactic, language ability, and individuals with ID are known to have difficulties across these domains of language, measures of these areas would have added to the findings of the current investigation. Furthermore, measures of other psycholololinguistic domains (e.g., working memory) and intelligence would be useful in the future.

The final limitation of the study was the number of students who participated. After inclusion criteria had been met, 39 students ultimately participated. More students would have led to greater power and certainty in the findings seen across the sample. Nonetheless, the size of the sample was sufficient to detect important differences and add to our understanding of questions for individuals with ID.

Directions for Future Research

In addition to addressing the limitations as described, there are several important areas that should be explored related to the findings of the current study. First, there are several implications of the current study in relation to intervention. Although there were not

significant differences found between the picture condition and no picture condition, this may have been due to some of the methodological issues described above. It was clear from anecdotal evidence and the topic-related responses that participants knew they were supposed to use the picture in some way to answer the question; however, it seems as though students need to be taught how to use pictures more effectively. Given that previous research has indicated that scaffolding procedures can work with students with ID to help them understand more conceptually complex questions raised about a text (Zetlin & Gallimore, 1983), it seems that the same procedures would be useful in teaching students with ID to use pictures as a scaffold to answer conceptually complex questions. However, the findings of the current study suggest that perhaps picture referents are not required as an intermediary step and that students with ID may do just as well with interventions that target answering conceptually complex questions without additional support.

The impact of the picture referent is one area to explore further in future research, but the impact of the conceptual level of the questions also warrants further investigation. The results of the current study and previous research indicate that students with ID may have more difficulty understanding higher-level questions than lower-level questions (Hewitt, 1998). As higher–level questions may facilitate greater understanding of material, and it has been shown that students with ID can be taught how to understand higher-level questions with scaffolding techniques (Zetlin & Gallimore, 1983), it is clear that the educational implications of these types of questions for these students should be explored further.

There are also interesting implications in terms of investigating question comprehension and answering in students who require augmentative and alternative communication (AAC). The AAC literature is replete with examples that point to the

relationship between question answering and communication and literacy skill development. Within discourse contexts such as conversation and storybook reading, conversation partners typically use closed, yes/no questions when interacting with individuals who use AAC (Light, Collier, & Parnes, 1985; Light & Kelford-Smith, 1994). The use of open-ended wh-questions is suggested for communication partners of individuals who use AAC in order to facilitate interaction in both conversation (Kent-Walsh & McNaughton, 2005) and storybook reading (Binger, Berens, Kent-Walsh, & Taylor, 2008; Kent-Walsh, Binger, & Hasham, 2010). The use of these types of questions has also been shown to be related to increased production of semantic concepts in individuals who use AAC (Kent-Walsh et al., 2010) and may also be important to syntactic development (Binger et al., 2008). Although the importance of comprehension of wh- questions is clear, no study has systematically investigated this in individuals who use AAC.

Conclusions

The current study was an investigation of comprehension of questions with six basic wh- question words in individuals with ID. The effect of a picture referent and no referent on comprehension was also investigated. Additionally, the relationship between receptive vocabulary ability and success in answering questions as a whole and at different conceptual levels was explored. Errors were also analyzed in order to determine if there were patterns that existed in the way students answered questions when they did not understand the questions.

The results of the study indicated that there were differences in the probability of answering different types of wh- questions. There were differences in the statistical probability of answering concrete (i.e., *who*, *what*, *where*) versus abstract questions (e.g., *when*, *why*, *how*) questions. Additionally, receptive vocabulary ability as measured by the

PPVT-4 (Dunn & Dunn, 2007) was found to be positively related to the probability of answering questions correctly. Close inspection of the relationship between receptive vocabulary and answering concrete and abstract questions revealed that receptive vocabulary was differentially related to the ability to answer concrete versus abstract questions. That is, there was less of a difference in the modeled probability of answering questions correctly for all of the participants while there was a more extreme difference in the modeled probability based on whether students had, low, average, or high receptive vocabulary relative to the population sample. There was not an overall statistically significant difference found between the picture and no referent condition.

Further investigation of errors revealed that there may have been some important differences in the incorrect answers students made based on whether they were asked in the picture condition or the no-picture condition. Further, the vast number of responses that were categorically related to *when* and *why* questions indicates that there may be an intermediate stage in comprehension in which some participants have knowledge about a particular question word, but do not understand enough to provide an accurate answer.

There are many implications of these results in terms of question knowledge and intervention. The fact that there are differences between concrete and abstract questions is important in that teachers and speech-language pathologists are often charged with helping these students understand all levels of questions. Further, the understanding of the decontextualized language often used in classrooms and necessary to comprehend text is often tied to the same type of language needed to understand more abstract types of questions. That is, students need to learn to answer questions and make connections that go beyond the here and now in order to successfully understand academic discourse.

The results of the current study highlight the need to better understand question answering in this population and develop better ways to teach students with ID how to answer questions. This study provides a base from which this type of research can continue. Question asking and answering exchanges are basic components of communication, literacy, and academic success. This needs to be addressed in greater detail with individuals with ID to ensure they have similar opportunities for growth as individuals who are typically developing.

.

APPENDIX A

Response Scoring Manual

General rules:

- No penalizing for grammar, but must be able to determine the student's general idea
- Pointing at the picture or somewhere in the room is acceptable as long as the notes indicate where the participant was that the participant's referent is clear

Who – No Picture

- Can refer to any logical person or group of people
- Can refer to a group of people in the class (e.g., "my class")
- Can also refer to people outside of class (e.g., "my father", "my family")

Who – Picture

- Answer can be "the girl"
- Reference to the boy's classmates
- It cannot be a random person such as "my mom" or "friends"

What – No Picture

- Any item of food
- Groups of food (e.g., fast food)
- Answer can be all inclusive of food (e.g., "I like all food.")

What – Picture

- Has to be an item of food in the picture
- Biscuit, chicken nuggets, cake, mac and cheese, cookie, cheese, potatoes, french fries, cheetos, bread
- Cannot be the girl's food
- If the student names food that is not in the picture, it is wrong
- Can be an appropriate group of food – e.g., "Junk food"

Where – No Picture

- Has to mention a place where lunch could traditionally be eaten
- Cafeteria, in the cafeteria, at the table, in school – all appropriate
- Home, restaurant appropriate as well
- Right here (the student could very well eat in the classroom)
- It is okay for students to incorrectly use the prepositions "in" or "at" when describing the place

Where – Picture

- Right there – if it is clear that the student points to where the students are eating in the picture
- At the table
- In the cafeteria
- School

When all

- A time is acceptable as long as it is an accepted time for lunch (between 10 and 12)
- Afternoon (but not morning)
- Can be a causal "*when*" time – (e.g., "*when* my teacher says", "*when* it's time to eat"); as long as it is logical (e.g., "*when* I get ready" would not work)
- lunchtime
- Can give credit for responses that indicates how long they have until lunch, if it is very clear, if not incorrect should be marked because the response could be about how long lunch is (e.g., "I have lunch in four hours" is acceptable while "four hours" is not)
- If a student mentions a certain time or period during the school day relative to other events (e.g., "before reading"), that is acceptable as there is no way to tell if that is a correct answer or not; but something like "before" by itself is not
- Needs to be relatively specific in regards to "*when*" – here, "anytime" would not work because it could be right or it could be the same thing as saying that lunch is at 7:30
- FOR PICTURE ONLY- An answer involving "now" is okay because it is more specific in terms of when lunch is in the setting of the picture, but this would not work for the no picture condition

Why all – Remember although grammar does not matter, it must be logical

- Should include a causal word in the response (e.g., so, to, because) (e.g., there are exceptions here, if it is makes sense as a response to a *why* question and the causal word is omitted and it still makes sense – e.g., "He doesn't want to be hungry" – this usually works just with because)
- Must be a logical response
- Can include a reference to food or a state of hunger (e.g., so I don't starve, because I'm hungry)
- Can also contain reference to pleasure or enjoyment gained from eating lunch("because I like it")
- Can contain a reference to food (e.g., "because I like food")
- Can also be answer about being told to eat (e.g., "because my mother tells me to.")
- NO – A response with a causal word that does not make a direct connection to the situation (e.g., "because he sits with his teacher"; "because I say thank you"; "because it's what I do")

How – All

- A response related to the process of how you would get lunch
- These answers can be a one-word response
- E.g., "pay" this is very definitely referring to the act of buying lunch – any reference to that transaction (going in line, paying with money, etc.) counts
- It should also be kept in mind that several of the students are on free or reduced lunch
- Additionally, naming a person/s who would help you get the lunch – this needs to be fairly specific in some manner (e.g., it has to refer to someone who would help with the act of getting lunch like a teacher or a person who works at the cafeteria)
- "lady" should be interpreted in this context as "lunch lady" (it is unlikely that students would use this word to refer to anyone else)
- Naming the place where they go to get the lunch also works (e.g., "the cafeteria", "from the cafeteria")

APPENDIX B

Error Analysis

There are six categories for incorrect answers:

1. Categorically related
2. Substitutions
3. I don't know/No response
4. Repetitions
5. Topic related
6. Unrelated

1. Categorically related
    - Responses to ***who*** questions are counted as categorically related if they include reference to a person or a group of persons
    - Responses to ***what*** questions are considered categorically related if they include any food item or group of foods
    - Responses to ***where*** questions are considered categorically related if they include a a response to a place
    - Responses to ***when*** questions are counted as categorically related when they include some aspect of time. For example:
        - 9:00
        - Like 8 hours
        - Causal-temporal responses (e.g., "*When* I bought toys")
    - Responses to ***why*** questions are counted as categorically related when the include any causal word, but do not answer the question – this is even the case if it combines a causal word and then talks about something in the picture, but does not answer the "*why*" question. For example:
        - "because I like toys"
        - "so I could go to recess"
    - Response to ***how*** questions are counted as categorically related when they provide information that implies the manner (e.g., quickly) or means (e.g., "he went to the store") in which something is done – although it does not necessarily apply to how someone would get lunch.
        - Quickly
        - "You go out to the playground."
        - A person that is not specifically identified as someone who would give a student lunch

2. Substitutions
    - If a different question word could be substituted for the one that was asked and the answer would be <u>correct</u> for the resulting question the response is counted as a repetition. For example:
        - If a student responded to "*Why* does the boy eat lunch?" with "5:30," it would not be counted as a substitution because it is an unacceptable answer for the *when* picture condition question, "*When* does the boy eat lunch?"
        - If a student responded to "*When* does the boy eat lunch? With "in the cafeteria" it would be counted as a substitution because the response is an excepted response to the substituted question, "*Where* does the boy eat lunch?"
    - Substitutions are allowed in the following questions (these are about the picture, but they correspond to the no picture condition):
        - *Who* does the boy eat with? – *what*
        - *Where* does the boy eat lunch? – *when, where, why, how*
        - *When* does the boy eat lunch? – *when, where, why, how*
        - *Why* does the boy eat lunch? – *when, where, why, how*
        - *How* does the boy get lunch? – *when, where, why, how*
    - All of the questions used in the study are at the end of this document

3. I don't know/No response
    - All responses that clearly "I don't know" are counted here.
    - All responses that are marked as "No response" are counted here.

4. Repetition
    - Responses are counted as repetitions if the response is comprised solely of words that appeared in the question. For example,
        - In response to the question "*Where* does the boy eat lunch?" the response is "lunch"
    - If there are other words in the response (even if they are not counted as part of the answer), these will not be counted as a repetition. For example,
        - In response to the question, "What food does the boy like?" the student replies, "The cafeteria. I don't know. Boy." The inclusion of "Boy" would be the answer, but it would not be a repetition.
5. Topic related
    - Picture condition - The response does not meet any of the 4 categories above but does pertain to the picture. For example:
        - "cookie",

        - "He is eating"
    - The response must appear in the picture so responses that are not directly in the picture are <u>not</u> coded as Topic Related. For example,
        - "He went to recess."
    - The response does not meet any of the 4 categories above but does include information <u>related to lunch</u> in some way. For example:
        - Any food item
        - Anything pertaining to lunch or eating that can be directly inferred (for example, incorrect times or substitutions that would make it to this level of analysis would not be counted here unless you could tell they specifically had to do with lunch).
6. Unrelated
    - All responses that do not fit into the other categories are counted as unrelated.

Questions used in study:

Picture	**No Picture**
Who does the boy eat with?	*Who* do you eat with?
What food does the boy like?	*What* food do you like?
Where does the boy eat lunch?	*Where* do you eat lunch?
When does the boy eat lunch?	*When* do you eat lunch?
Why does the boy eat lunch?	*Why* do you eat lunch?
How does the boy get lunch?	*How* do you get lunch?

APPENDIX C

Sample Question Comprehension Form

Picture Questions

Say: "I'm going to ask you some questions about a picture. In this picture there is a boy and a girl eating lunch (point to each person in the picture). I want you to listen carefully. I want you to do your best to tell me the answer to the questions I ask. Ready?"

Note: Repetitions are allowed if the student did not hear the question.

Write: Write what the child says verbatim in the space provided. Also, write any additional notes in that area.

1. *When* does the boy eat lunch?	
2. *What* food does the boy like?	
3. *How* does the boy get lunch?	
4. *Who* does the boy eat with?	
5. *Why* does the boy eat lunch?	
6. *Where* does the boy eat lunch?	

Lunch Questions

Say: "I'm going to ask you some questions about lunch. I want you to listen carefully. I want you to do your best to tell me the answer to the questions I ask. Ready?"

Note: Repetitions are allowed if the student did not hear the question.

1. *Where* do you eat lunch?	
2. *What* food do you like?	
3. *When* do you eat lunch?	
4. *How* do you get lunch?	
5. *Who* do you eat with?	
6. *Why* do you eat lunch?	

REFERENCES

Abbeduto, L., Brady, N., & Kover, S. T. (2007). Language development and fragile X syndrome: profiles, syndrome-specificity, and within-syndrome differences. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(1) 36-46.

Abbeduto, L., Furman, L., & Davies, B. (1989). Relation between the receptive language and mental age of persons with mental retardation. *American Journal on Mental Retardation, 93*, 535-543

Abbeduto, L., Murphy, M. M., Cawthon, S. W., Richmond, E. K., Weissman, M. D., Karadottir, S., & O'Brien, A. (2003). Receptive language skills of adolescents and young adults with down or fragile X syndrome. *American Journal of Mental Retardation, 108* (3), 149-160.

Abbeduto, L., Pavetto, M., Kesin, E., Weissman, M., Karadottir, S., O'Brien, A., & Cawthon, S. (2001). The linguistic and cognitive profile of Down syndrome: Evidence from a comparison with fragile X syndrome. *Down Syndrome Research and Practice, 7*(1), 9-15.

Abbeduto, L., Warren, S. F., & Conners, F. A. (2007). Language development in Down syndrome: From the prelinguistic period to the acquisition of literacy. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(3), 247-261.

Binger, C., Berens, J., Kent-Walsh, J., & Taylor, S. (2008). The effects of aided AAC interventions on AAC use, speech, and symbolic gestures. *Seminars in Speech and Language, 29*(2), 101-11.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley.

Bloom, L., Merkin, S., & Wootten, J. (1982). "Wh"-Questions: Linguistic Factors That Contribute to the Sequence of Acquisition. *Child Development, 53*(4), 1084-1092.

Brown, R. (1968). The development of wh questions in child speech. *Journal of Verbal Learning and Behavior, 7*, 279-290.

Cairns, H. S., & Hsu, J. R. (1978). Who, why, when and how: A developmental study. *Journal of Child Language, 5*, 478-488.

Chapman R.S. (2006). Language learning in Down syndrome: The speech and language profile compared to adolescents with cognitive impairment of unknown origin. *Down Syndrome Research and Practice,10,* 61–66.

Chapman, R. S., Hesketh, L. J., & Kistler, D. J. (2002). Predicting Longitudinal Change in Language Production and Comprehension in Individuals with Down Syndrome:

Hierarchical Linear Modeling. *Journal of Speech, Language, and Hearing Research, 45,5,* 902-15.

Chapman, R. S., Seung, H.-K., Schwartz, S. E., & Bird, E. K.R. (1998). Language Skills of Children and Adolescents with Down Syndrome: II. Production Deficits. *Journal of Speech, Language, and Hearing Research, 41,* 4, 861-73.

Clancy, P. M. (1989). Form and Function in the Acquisition of Korean Wh-Questions. *Journal of Child Language, 16* (2), 323-47.

Common Core State Standards Initiative (2010). Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects: Reading Standards for Literature K-5. http://www.corestandards.org.

Deevy, P., & Leonard, L. B. (2004). The comprehension of wh-questions in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 47*(4), 802.

de Rivera, C., Weitzman, E., Greenberg, J., & Girolametto, L. (2005). Children's responses to educators' questions in day care playgroups. *American Journal of Speech-Language Pathology, 14,* 14-26.

De, V. J., Roeper, T., Pearson, B., & Bland-Stewart, L. (2008). Answering hard questions: Wh-movement across dialects and disorder. *Applied Psycholinguistics, 29,* 1, 67-103.

Dillon, J. T. (1988). Questioning and teaching: A manual of practice. New York: Teachers College.

Dunn, L.M. (1965). The Peabody Picture Vocabulary Test. Circle Pines, MN: American Guidance Service.

Dunn & Dunn (2007). Peabody Picture Vocabulary Test – 4th. Pearson.

Ervin-Tripp, S. (1970). Discourse agreement: How children answer questions. In R. Hayes (Ed.), *Cognition and language learning* (pp. 79-107). New York: Wiley and Sons.

Facon B, Facon-Bollengier T, Grubar JC. 2002. Chronological age, receptive vocabulary, and
syntax comprehension in children and adolescents with mental retardation. *American Journal on Mental Retardation, 107,* 91–98.

Fidler, D., Philofsky, A., & Hepburn, S. (2007). Language phenotypes and intervention planning: Bridging research and practice. *Mental retardation and developmental disabilities research reviews, 13*(1), 47-57.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *Journal of Autism and Developmental Disorders, 33,* 4, 365-82.

Fowler, A. E. (1998). Language in mental retardation: Associations with and dissociations from general cognition. (E. Zigler, with J. A. Burack, Robert M Hodapp, Eds.), *Handbook of mental retardation and development*, 290-333. Cambridge University Press.

Friedmann, N, & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of wh questions in three subtypes of SLI. *Lingua, 121*(3), 367-382.

Hewitt, L. E. (1998). Influence of question type on response adequacy in young adults with autism. *Journal of Communication Disorders, 31*(2), 135-152.

Hodapp, R., Griffin, M., Burke, M., & Fisher, M. (2011). Intellectual disabilities. In Steinberg, R. & Kaufman, S. (Eds.), *The Cambridge book of intelligence* (pp. 193-209). New York: Cambridge University Press.

Hundert, J., & van Delft, S. (April 29, 2009). Teaching Children With Autism Spectrum Disorders to Answer Inferential "*Why*" Questions. *Focus on Autism and Other Developmental Disabilities*, 24, 2, 67-76.

Hunkins, F. P. (1989). *Teaching thinking through effective questioning*. Boston: Christopher-Gordon.

Jahr, E. (2001). Teaching children with autism to answer novel wh-questions by utilizing a multiple exemplar strategy. *Research in Developmental Disabilities, 22*(5), 407-423.

Joffe, V., & Varlokosta, S. (2007). Patterns of syntactic development in children with Williams syndrome and Down's syndrome: Evidence from passives and wh questions. *Clinical Linguistics & Phonetics, 21*(9), 705-727.

Keller-Bell, Y. D., & Abbeduto, L. (2007). Narrative development in adolescents and young adults with fragile X syndrome. *American Journal on Mental Retardation, 112*, 289-299.

Kent-Walsh, J., Binger, C., & Hasham, Z. (2010). Effects of parent instruction on the symbolic communication of children using AAC during storybook reading. *American Journal of Speech-Language Pathology, 19*, 97-107.

Kent-Walsh, J., & Mcnaughton, D. (2005). Communication Partner Instruction in AAC: Present Practices and Future Directions. *Augmentative & Alternative Communication, 21*, 3, 195-204.

Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Topics in Language Disorders, 25*(1), 51-64.

Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16*, 287–308.

Krantz, P., Zalenski, S., Hall, L., Fenske, E., McClannahan, L. (1981). Teaching complex language to autistic children. *Analysis and Intervention in Developmental Disabilities, 1*, 259-297.

Laws, G., & Bishop D.V.M. (2003). A comparison of language abilities in adolescents with Down syndrome and children with specific language impairment. *Journal of Speech Language Hearing Research, 46*, 1324–1339.

Lee, R. F., & Ashmore, L. L. (1983). Receptive and expressive wh-question performance by language-delayed children. *Journal of Communication Disorders, 16*(2), 99-109.

Lehnert, W. G. (1978). *The Process of Question Answering*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Leinonen, E., Letts, C., & Smith, B. R. (2000). *Children's pragmatic communication difficulties*. Whurr Publishers, London.

Light, J., Collier, B., & Parnes, P. (1985). Communicative interaction between young nonspeaking physically disabled children and their primary caregivers: Part I – Discourse patterns. *Augmentative and Alternative Communication, 1*, 74 – 83.

Light, J., & Kelford Smith, A. (1993). Home literacy experiences of preschoolers who use AAC systems and of their nondisabled peers. *Augmentative and Alternative Communication, 9*, 10–25.

Loukusa, S., Ryder, N., & Leinonen, E. (2008). Answering Questions and Explaining Answers: A Study of Finnish-Speaking Children. *Journal of Psycholinguistic Research, 37*(3), 219-241.

Massey, S. L., Pence, K. L., Justice, L. M., & Bowles, R. P. (2008). Educators' use of cognitively challenging questions in economically disadvantaged preschool classroom contexts. *Early Education & Development, 19*(2), 340-360.

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities, 30*(6), 1107-1114.

McDuffie A, Abbeduto L. (2009). Developmental delay and genetic syndromes: Down syndrome, fragile X syndrome, and Williams syndrome. In: Schwartz R, editor. *Handbook of child language disorders*. Psychology Press; New York.

Mervis, C. B., & Becerra, A. M. (2007). Language and communicative development in Williams syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(1), 3-15.

Morgan, M. F., Moni, K. B., & Jobling, A. (2009). Who? Where? What? When? Why? How? Question words - What do they mean? *British Journal of Learning Disabilities, 37*(3), 178-185.

National Institue of child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

O'Grady, W. D. (1997). *Syntactic development*. Chicago: University of Chicago.

Owens, R. E. (2008). *Language development: An introduction* (7th ed.). Needham Heights, MA: Allyn & Bacon.

Parnell, M. M., & Amerman, J. D. (1983). Answers to wh-questions: research and application. In T. Gallagher & C. Prutting (Eds.), *Pragmatic assessment and intervention issues in language*. San Diego, CA. College-Hill Press.

Parnell, M. M., Amerman, J. D., & Harting, R. D. (1986). Responses of language-disordered children to wh-questions. *Language, Speech, and Hearing Services in Schools, 17*(2), 95-106.

Parnell, M. M., Patterson, S. S., & Harding, M. A. (1984). Answers to wh-questions: A developmental study. *Journal of Speech and Hearing Research, 27*(2), 297-305.

Paul, R., Dykens, E., Leckman, J. F., Watson, M., Breg, W. R., & Cohen, D. J. (1987). A comparison of language characteristics of mentally retarded adults with fragile X syndrome and those with nonspecific mental retardation and autism. *Journal of Autism and Developmental Disorders, 17*, 4, 457-68.

Paul, R. (2007). *Language disorders from infancy through adolescence: Assessment & intervention*. Edinburgh: Elsevier Mosby.

Price, J., Roberts, J., Vandergrift, N., & Martin, G. (2007). Language comprehension in boys with fragile X syndrome and boys with Down syndrome. *Journal of Intellectual Disability Research, 54*(4), 318-326.

Radford, A. (2009). *Analysing English sentences: A minimalist approach*. Cambridge, UK: Cambridge University Press.

Roberts, J. E., Hennon, E. A., Price, J. R., Dear, E., Anderson, K., & Vandergrift, N. A. (2007). Expressive Language during Conversational Speech in Boys with Fragile X Syndrome. *American Journal on Mental Retardation, 112*, 1, 1-17.

Roberts, J., Martin, G. E., Moskowitz, L., Harris, A. A., Foreman, J., & Nelson, L. (2007). Discourse skills of boys with fragile X syndrome in comparison to boys with Down syndrome. *Journal of Speech, Language, and Hearing Research , 50*(2), 475-92.

Roberts, J., Price, J., & Malkin, C. (2007). Language and communication development in down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(1), 26-35.

Roberts, J., Price, J., Barnes, E., Nelson, L., Burchinal, M., Hennon, E. A., Moskowitz, L., Hooper, S. R. (2007). Receptive Vocabulary, Expressive Vocabulary, and Speech Production of Boys with Fragile X Syndrome in Comparison to Boys with Down Syndrome. *American Journal on Mental Retardation, 112*(3), 1-17.

Roeper, T. (January 01, 2004). Diagnosing language variations: underlying principles for syntactic assessment. *Seminars in Speech and Language, 25*(1), 41-55.

Rosenberg, S., & Abbeduto, L. (1993). *Language and communication in mental retardation: Development, processes, and intervention*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language, 30,* 609–635.

Secan, K. E., Egel, A. L., & Tilley, C. S. (1989). Acquisition, generalization, and maintenance of question-answering skills in autistic children. *Journal of Applied Behavior Analysis, 22*(2), 181-96.

Seidl, A., Hollich, G., & Jusczyk, P. (2003). Early Understanding of Subject and Object Wh Questions. *Infancy, 4*(3), 423-436.

Shevell, M. (2008). Global developmental delay and mental retardation or intellectual disability: conceptualization, evaluation, and etiology. *Pediatr Clin North Am, 55*(5), 1071-1084.

Shulman, B. B., & Capone, N. C. (2009). *Language Development: Foundations, Processes and Clinical Applications*. Boston, MA: Jones and Bartlett Learning.

Stromswold, K. (1995). The acquisition of subject and object wh-questions. *Language Acquisition, 4*(1), 5-48.

Tager-Flusberg, H. (2005). Putting words together: morphology and syntax in the preschool years. In Berko-Gleason, J. (Ed.) *The development of language (6th Ed).* Boston: Pearson.

Thorndike, E. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words.* New York: Teachers College Columbia University.

Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language, 4*(2), 211-224.

U.S. Department of Education (2002). *Twenty-fourth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, D.C.

van der Schuit, M., Segers, E., van Balkom, H., & Verhoeven, L. (2011). How cognitive factors affect language development in children with intellectual disabilities. *Research in Developmental Disabilities, 32*(5), 1884-1894.

Vicari, S., Caselli, M.C. & Tonucci, F. (2000). Early language development in Italian children with Down syndrome: asynchrony of lexical and morphosyntactic abilities. *Neuropsychologia*, 38, 634-644.

Winzemer, J. A. (1981). *A Lexical-expectation Model for Children's Comprehension of Wh-questions*. Unpublished doctoral dissertation, City University of New York.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

Zetlin, A. G., & Gallimore, R. (1980). A Cognitive Skills Training Program for Moderately Retarded Learners. *Education and Training of the Mentally Retarded, 15*(2), 121-131.

Zetlin, A. G., & Gallimore, R. (1983). The Development of Comprehension Strategies through the Regulatory Function of Teacher Questions. *Education and Training of the Mentally Retarded, 18,* 3, 176-84.