Georgia State University
ScholarWorks @ Georgia State University

Psychology Theses | Department of Psychology

5-2-2018

Patterns of Communication, Cognition, and Adaptive Behavior in Children with Developmental Delays

Phebe Albert

Follow this and additional works at: https://scholarworks.gsu.edu/psych_theses

Recommended Citation
Albert, Phebe, "Patterns of Communication, Cognition, and Adaptive Behavior in Children with Developmental Delays." Thesis, Georgia State University, 2018.
doi: https://doi.org/10.57709/11906915

This Thesis is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

PATTERNS OF COMMUNICATION, COGNITION AND ADAPTIVE BEHAVIOR IN CHILDREN WITH DEVELOPMENTAL DELAYS

by

PHEBE ALBERT

Under the Direction of MaryAnn Romski, PhD

ABSTRACT

Young children with developmental disabilities (DD) can demonstrate a wide range of difficulties in different domains including cognition, language and adaptive behaviors. Accurately assessing these difficulties and characterizing patterns of strengths and weaknesses is important for informing intervention strategies (Ben-Sasson & Gill, 2014; Plesa Skwerer, Jordan, Brukilacchio & Tager-Flusberg, 2016). The current study examines how toddlers with a significant developmental delay and less than 10 spoken words perform across different developmental domains, (i.e., cognitive, language and adaptive functioning) and across assessment methods, (i.e., parent report and clinician-administered). Results indicated that parent-reported and clinician-administered measures of cognition, language and adaptive functioning are highly related, as are young children's performances across these domains. Findings also revealed that children with similarly limited spoken language can exhibit a variety of strengths and weaknesses in other domains.

INDEX WORDS: Developmental disabilities, Language, Cognition, Adaptive behavior

PATTERNS OF COMMUNICATION, COGNITION AND ADAPTIVE BEHAVIOR IN CHILDREN WITH DEVELOPMENTAL DEALYS

by

PHEBE ALBERT

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2018

Copyright by
Phebe Elizabeth Albert
2018

PATTERNS OF COMMUNICATION, COGNITION AND ADAPTIVE BEHAVIOR ABILITIES IN CHILDREN WITH DEVELOPMENTAL DELAYS

by

PHEBE ALBERT

Committee Chair: MaryAnn Romski

Committee: Rose A. Sevcik

Robin D. Morris

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2018

ACKNOWLEDGEMENTS

I would like to thank my committee members, Dr. Rose Sevcik and Dr. Robin Morris, and specially my academic advisor and committee chair, Dr. MaryAnn Romski, for all of their guidance and support throughout this project. The studies utilized in this project were supported by grant DC-03799 from the National Institute on Deafness and Other Communication Disorders and grant R324A070122 from the U.S. Department of Education Institute of Education Sciences to Dr. MaryAnn Romski.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Assessing children with developmental delays

Young children with developmental delays (DD) demonstrate a wide range of difficulties across different domains of functioning. Skill deficits persist throughout later development, and these children often experience impairments across multiple developmental domains including cognition, motor skills, adaptive behaviors and communication (Shevell, Majnemer, Platt, Webster, & Birnbaum, 2005). The characterization and accurate assessment of these deficits is important for informing intervention strategies (Ben-Sasson & Gill, 2014; Plesa Skwerer et al., 2016); it is well documented that early intervention has positive effects for developmental outcomes in children with developmental delays (American Speech-Hearing-Language Association [ASHA], 2008; (Buschmann, Multhauf, Hasselhorn, & Pietz, 2015; Ciccone, Hennessey, & Stokes, 2012). Psychologists and other health professionals rely on a variety of assessment methods to measure the abilities of children with DD, and the growth of these abilities over time, including direct observation, parent interview and questionnaires. The current study will examine how toddlers with a significant developmental delay and less than 10 spoken words perform across several developmental domains on a series of assessment measures that include direct testing and parent report.

1.1.1 Challenges of assessing toddlers with developmental delays

A number of measures have been developed to assess young children's development in cognition, adaptive behavior, and language and communication. Many of these measures also are utilized to assess children with developmental delays. However, obtaining reliable and valid assessments of abilities in young children with developmental delays remains a complex task for clinicians and researchers. A myriad of challenges arise when professionals across fields employ standardized measures to assess the developmental status of young children with developmental disabilities, such as questionable technical adequacy and the inextricable impact of deficits in one domain, such as motor skills, on the ability to perform on measures of other domains, such as nonverbal cognition (Bradley-Johnson, 2001; DeVeney, Hoffman, & Cress, 2012). Also, the interpretation and integration of research findings across studies can be impacted by differences in how children are assessed; including which assessments and which versions of assessments are used. Some of the developmental abilities of interest in this study, such as language and adaptive functioning, are measured differently across tests (Magiati & Howlin, 2001), and developmental assessments continuously undergo revisions and updates that limit the comparisons that can be made between different versions of the same assessment used across studies.

Another challenge of assessing young children with developmental disabilities is that their performance on assessment measures is often hindered by problem behaviors that are common among this population (Baker, Blacher, Crnic, & Edelbrock, 2002; Hauser-Cram & Woodman, 2016). High rates of challenging behaviors, including aggression, tantrums, hyperactivity, defiance and reduced engagement in play behaviors have been documented in toddlers with general developmental delays due to genetic syndromes (e.g., down syndrome) and unknown etiologies (Keller & Fox, 2009; Krakow & Kopp, 1983). Assessment performance is also impacted by other factors associated with testing such as difficulties gaining and maintaining attention throughout testing, environmental distractions such as discomfort associated with the unfamiliarity of the testing setting and/or frustration, anxiety and difficulty comprehending test instructions (Tager-Flusberg, 2000). For example, Akshoomoff (2006) investigated performance on the Mullen Scales of Early Learning (Mullen; Mullen, 1995) in a sample of toddlers, ages 16 to 43 months who met criteria for autism spectrum disorder (ASD) according to autism specific parent interview and direct testing (i.e., the Autism Diagnostic Interview and the Autism Diagnostic Interview Schedule), and who did not have any other comorbid medical conditions, preterm birth, hearing, motor or visual impairments. The children in this sample demonstrated significantly lower scores across all Mullen developmental domains (Receptive Language, Expressive Language, Fine Motor and Visual Reception) than typically developing toddlers, and 76% of the ASD sample performed at the lowest ability level on one or more Mullen domains, represented by a T-score of 20 or <20. Akshoomoff (2006) found that Mullen performance in this sample of children with ASD was negatively correlated with being off-task during testing and positively correlated with level of engagement during testing.

Much of the current literature on assessment and profiles of developmental abilities in toddlers with developmental disabilities focuses on children with ASD (Akshoomoff, 2006; Bruckner, Yoder, Stone, & Saylor, 2007; Magiati & Howlin, 2001; Nordahl-Hansen, Kaale, & Ulvund, 2014; Plesa Skwerer et al., 2016). Often times research that does include children with other developmental delays focuses primarily on group differences between those children and children with ASD (Fodstad, Matson, Hess, & Neal, 2009; Provost, Lopez, & Heimerl, 2007). While young children with ASD share common deficits in communication and language abilities with children who have general developmental delays, their ability profiles differ (Čeponienė et al., 2003; Fodstad et al., 2009; Swineford, Guthrie, & Thurm, 2015; Weismer, Lord, & Esler, 2010; Wilkins & Matson, 2009). Interestingly, Seol et al. (2014) compared performance on a Korean language scale, the Sequenced Language Scale for Infants (SELSI), between a group of children with ASD and a group of children with developmental language delay (DLD) and found that differences in language profiles between children with ASD and those with DLD were even more distinct at younger ages, i.e., from 20-29 months than from 40-49 months. Overall, these findings underscore the importance of broadening the current literature on developmental assessment performance in young children to include those with general developmental delays.

1.1.2 Challenges related to assessing very young children

The appropriateness of cognitive assessment in young children has been a controversial topic amongst researchers and clinicians because of difficulties defining and operationalizing the construct of intelligence in very young children, which also contributes to difficulty establishing the predictive validity of assessment performance in this population (Bradley-Johnson, 2001). These challenges and assessment inadequacies are especially amplified in children under the age of three. Bradley-Johnson (2001) conducted a review of the technical adequacy of six developmental assessments and presented findings on their reliability, validity, item gradients, floors and ceilings, and standardization procedures. She found that common problems across tests included inadequate floors, which inhibited their ability to diagnose intellectual disability, outdated normative samples and a lack of normative data reported by age level. The ability of a single developmental assessment such as the Bayley Scales of Infant Development, 2nd edition (BSID-II; Bayley, 1993) and/or the Battelle Developmental Inventory (BDI; Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1987) to validly identify developmental delay is also reduced in young children under the age of three (Gerken, Eliason, & Arthur, 1994). However, developmental assessment can make meaningful contributions to intervention planning and service eligibility when these issues are taken into consideration and a multiple-measure assessment approach with routine follow-ups is utilized (Bradley-Johnson, 2001).

Another important consideration of assessment in young children with developmental disabilities is the use of raw scores, standard scores and/or age equivalent (AE) scores to measure ability level. These scores are derived differently and are not interchangeable (Sattler & Hoge, 2006; Sullivan, Winter, Sass, & Svenkerud, 2014). Standard scores provide the most psychometrically robust information, they are comparable across different assessment measures and they provide information about an individual's range of performance compared to a normative sample of peers (Sullivan et al., 2014). However, because standard scores are derived from comparison to typical performance, they may mask small or atypical changes in performance over time. Thus, for children who demonstrate atypical development, such as those with developmental disabilities, raw scores are often a more sensitive measure of change over time in individual performance. One of the most problematic characteristics of raw scores is that they are not comparable across measures (Sullivan et al., 2014). Some researchers also utilize age equivalent scores to describe the performance of children with developmental disabilities across developmental domains within a test (Weismer et al., 2010), and across different tests (Milne & McDonald, 2015). Carter et al. (1998) investigated profiles of performance across Vineland ABS domains in a sample of individuals with ASD, ages 2 to 59 years, with varying spoken language ability and found that the use of age equivalents produced the expected pattern of reduced socialization and communication skills compared to daily living skills, but the use of standard scores produced an unexpected pattern of performance according to previous research. Carter et al. posited that this discrepancy was at least in part caused by "range restriction due to basal effects" (p.299) when standard scores were employed in the profile analyses. They suggested that age equivalents might be a more useful measure of adaptive behavior performance profiles in groups of lower functioning individuals with limited spoken language. Milne and McDonald (2015) found that Vineland-II age equivalent scores in a sample of children, ages 23 months to 5 years with developmental disabilities produced a more accurate picture of developmental strengths and weaknesses, and demonstrated greater convergent validity with other measures of developmental level and adaptive functioning than did standard scores. They concluded that standard scores are better suited for identifying significant adaptive functioning impairment, while age equivalents might be better suited to identify areas and degree of needed support. However, on the Peabody Picture Vocabulary Test–Third Edition (PPVT-III; Dunn & Dunn, 1997) Sullivan, Winter, Sass and Svenkerud (2014) found that age equivalents from a sample of young children ages 3 to 5 years old demonstrated a significant floor effect and lacked the same precision as raw scores and standard scores, i.e., children with different raw scores and standard scores had the same age equivalent scores. This issue is particularly relevant for very young children and children with developmental disabilities who perform at the lower range of ability and are thus more vulnerable to floor effects (Sullivan et al., 2014). Alternatively, Akshoomoff (2006) utilized age equivalent scores instead of T-scores on the Mullen Scales of Early Learning to investigate relationships between developmental assessment performance and testing behaviors in a sample of toddlers with ASD. Akshoomoff chose to use age equivalents because a large number of participants with ASD received T-scores on MSEL subdomains of 20 or below, which is the lowest possible T-score. Therefore, in this case, age equivalents were able to represent a broader range of performance that more accurately captured the group of toddlers with ASD who were performing at the test floor. Taken together, these findings suggest that it is important to consider the psychometric properties of individual tests when making decisions about which scores will most accurately measure performance, especially in young children with developmental disabilities. Further, determining which score type to use or report depends largely on the research or referral questions that are being answered. For example, the assessment of developmental strengths and weaknesses of a young child with a developmental disability across measures and/or assessment domains may be best characterized by raw scores or age equivalents, whereas diagnostic decisions might best be addressed with standard scores that provide normative comparisons (Sullivan et al., 2014; Weismer et al., 2010).

1.1.3 Assessing toddlers with expressive language impairment

Developmental assessment of toddlers who have limited expressive language ability is especially challenging because many standardized tests require children to use spoken language to communicate test responses, which disadvantages young children who do not yet have spoken words (Brady, Marquis, Fleming, & McLean, 2004). Cirrin and Rowland (1985) found that individuals with severe intellectual disabilities, ages 10 to 18 years, utilized several preverbal communicative strategies such as signing, pointing to request an object, and physical contact or extending an object to request an action from a caregiver. They also found a large amount of individual difference across participants in the use and rate of preverbal communication. Thus, they suggested that a focus on traditional "[spoken] language behaviors" (p.60) may be excluding important preverbal communicative acts that can be utilized for Individualized Education Plan (IEP) development and treatment planning. Additionally, their findings suggest that grouping together all individuals who do not speak neglects important individual differences that are present even in lower functioning populations. Floor effects that occur on many standardized developmental assessments also can contribute to this misconception by creating an inaccurate picture of little to no differentiation between the functioning and abilities of individuals who do not have spoken language.

1.2 Assessment performance profiles in children with developmental delays

One way to evaluate the technical adequacy of standardized assessments in toddlers with developmental delays is to investigate patterns of performance both within and between testing measures. It is important to describe how this group of children performs on different domains (receptive language, fine motor skills, visual reception, etc.) within a measure to identify trends in individual and group performances across developmental areas. In addition to these patterns, it is important to compare domain scores across multiple measures, which can provide insight into measurement validity.

1.2.1 Performance across assessment domains

Previous research indicates that young children with developmental disabilities, ranging from 10 months to 6 years of age, demonstrate unique profiles of strengths and weaknesses across different domains of functioning (Caselli et al., 1998; Fidler, Philofsky, & Hepburn, 2007; Luyster, Seery, Talbott, & Helen, 2011; Singer Harris, Bellugi, Bates, Jones, & Rossen, 1997; Weismer et al., 2010). These profiles frequently contain incongruent performance across domains during the first few years of life. For example, Ben-Sasson and Gill (2014) found that in a community sample of toddlers with a mean age of 12 months (24% of the sample was diagnosed with ASD), increases in scores on one subdomain of the Mullen over time (approximately 17 months later) was frequently accompanied by decreases on scores in another domain. For instance, 61.1% of children who demonstrated a decrease over time on the Gross Motor domain exhibited an increase on at least one of the two language domains (Expressive and Receptive Language). Moreover, scores on the Mullen Early Learning Composite (Mullen ELC), which is used as a measure of overall developmental level that incorporates performance from all subdomains, did not show decreases over time in the group of children who demonstrated decreases over time in individual Mullen subdomain scores. These findings underscore the importance of examining subdomain performance when evaluating infants and toddlers. Overall developmental scores, such as the Mullen ELC, may dilute important information about strengths and weaknesses across different developmental domains that can play a role in support and intervention planning. For example, a child with impaired social communication skills who demonstrates developmentally appropriate motor skills might benefit from increased participation in playground activities with peers or bike riding with family members to improve social communication opportunities, while another child with similarly impaired social and cognitive abilities in addition to motor deficits might be unsafe during these same activities due to under-developed motor strength and coordination (Milne & McDonald, 2015).

1.2.2 Performance across measures

While it is well documented that profiles of strengths and weaknesses exist among children with developmental delays, it is still unclear whether they always reflect true differences in ability or are influenced by measurement issues. When young children with developmental disabilities are assessed with multiple measures, there are often discrepancies in their performances on the same domain across different tests (Magiati & Howlin, 2001; Scattone, Raggio, & May, 2011). For example, Plesa Skwerer et al. (2016) systematically compared multiple measures of receptive language ability in a sample of "minimally verbal" children and adolescents, 5 to 21 years old, with ASD. They found that receptive language performance varied significantly across assessment methods. These findings underscore the importance of evaluating multiple measures simultaneously when investigating developmental profiles of toddlers with developmental delays; one measure alone may not accurately reflect an individual's true abilities.

1.2.3 Parent report compared to direct assessment measures

Two of the most commonly utilized types of developmental assessment measures are direct assessments administered by trained clinicians and caregiver report measures. Direct assessment and parent and/or teacher report can both be valid tools for evaluating language, cognition and adaptive behavior in children. Each measurement type has unique strengths and weaknesses. Direct assessment provides a structured setting where trained professionals explicitly elicit skills and behaviors. Performance measured via direct testing benefits from standardization of how information is collected across individuals, and from interpretation by trained professionals. However, it can be impacted by a variety of factors including levels of child attention and engagement during testing, severity of cognitive impairment, and other factors related to the testing environment (Plesa Skwerer et al., 2016). Additionally, children with developmental disabilities, like ASD, may demonstrate different abilities within different contexts, which can contribute to discrepancies between the results of direct assessment and parent report of abilities (Nordahl-Hansen et al., 2014). Parent report measures can take into account typical performance across contexts and during daily life routines. However, they can be biased by inaccurate parent reporting, particularly a tendency to over-report comprehension ability in young children (Charman, 2004; Scattone et al., 2011). Relying solely on one type of assessment strategy can result in a child's abilities being over- or underestimated and inappropriate assignment to or rejection from intervention services (Scattone et al., 2011). Thus, it is recommended that a combination of multiple measures be used when assessing young children with developmental disabilities (Akshoomoff, 2006; Nordahl-Hansen et al., 2014).

Currently, research findings are mixed regarding the agreement between direct assessment and parent report in young children with developmental disabilities. Several studies have replicated strong correlations between some parent report and direct assessment indices of language and adaptive skills in toddlers, ages 12 to 42 months with developmental disabilities using the MacArthur Bates Communicative and Developmental Inventory (MCDI; Fenson,

2007) (Björn, Kakkuri, & Leppänen, 2014; Dale, 1991), the Vineland Adaptive Behavior Scales, 2nd Edition (VABS-II; Sparrow, Cicchetti, & Balla, 2005) (Scattone et al., 2011) and the Mullen and the Sequenced Inventory of Communication Development (SICD; Hendrick, Prather, & Tobin, 1984) (Weismer et al., 2010). Interestingly, while the majority of these studies have produced significant correlations between portions of parent report and direct assessment, the results are often inconsistent across test subdomains. For example, Björn et al. (2014) found that, in a sample of young children ages 12 to 18 months, there was a significant relationship between the number of words produced reported on the MCDI and expressive language scores on the Bayley Scales of Infant Development, 3rd edition (BSID-III; Bayley & Reuner, 2006), and between the number of words understood reported on the MCDI and receptive language scores on the Bayley-III. However, Bayley receptive language scores were not significantly correlated with other receptive language variables from the MCDI, i.e., *first signs of understanding instructions and questions*, *first signs of receptive language*, and *first communicative gestures*. Other studies also have found discrepancies between parent report and direct assessment agreement for expressive versus receptive language ability. Charman (2004) reported that test performance from a small subgroup of children, ages 18 months to 7 years with ASD who participated in a larger study (see Charman, Drew, Baird, & Baird, 2003), demonstrated a strong agreement between parent report on the MCDI and direct assessment of *language production*, N = 15, $r = .66$, $p < .01$, but a small non-significant correlation between parent report and direct assessment of *language comprehension,* N = 18, $r = .09$. Furthermore, Lyytinen, Laakso, Poikkeus, and Rita (1999) found that parent reports of toddler's *vocabulary production* at 14 months of age on the MCDI revealed greater predictive correlations with direct assessment of language and cognitive skills at 24 months than parent report of *vocabulary comprehension* at 14 months. Taken together, these findings reveal that the agreement between parent report and direct testing may not be consistent across different domains, and therefore it is important to assess multiple developmental domains when evaluating the agreement between parent report and direct assessment.

1.2.4 General Measures of Developmental Skills

Although evaluating the abilities of infants and toddlers with developmental disabilities is complex and often accompanied by many challenges, there are a number of measures that have been utilized to assess development abilities in this population.

1.2.4.1 The Mullen Scales of Early Learning

The Mullen Scales of Early Learning (Mullen) is a clinician-administered developmental assessment for children from birth to 68 months that measures performance in five independent domains: Gross Motor that assesses mobility and motor control, Visual Reception that assesses visual processing and memory, Fine Motor that assesses visual-motor ability, Receptive Language that assesses auditory comprehension and memory and Expressive Language that assesses speaking and language formation (Mullen, 1995).

The Mullen has demonstrated strong validity as a tool for assessing developmental strengths and weakness and overall developmental level in children with developmental disabilities (Swineford et al., 2015). Some research has noted that specific profiles of performance arise across Mullen domains in children, ages 16 to 43 months with ASD, i.e. relatively strong fine motor ability accompanied by relatively weak receptive language ability (Akshoomoff, 2006). Another study of children with autism spectrum disorder (mean age = 38 months), cerebral palsy (mean age = 39 months) and epilepsy (mean age = 28 months) revealed that while the Mullen was able to detect general developmental delays across this range of developmental disorders, it did not produce differential profiles of performance across diagnostic groups (Burns, King, & Spencer, 2013). However, research to date has not thoroughly examined multi-dimensional performance profiles in children with general developmental delays. This is surprising considering that children with developmental delays display deficits across a range of abilities, coupled with strong evidence that abilities in one domain of development, such as motor skills, influence abilities in other domains such as communication (Iverson, 2010). These findings underscore the importance of examining children's abilities across all developmental domains (Bradley-Johnson & Johnson, 2007).

1.2.4.2 The Vineland Adaptive Behavior Scales

The Vineland Adaptive Behavior Scales is one of the most commonly used measures of parent-reported adaptive behavior for individuals from birth through 90 years. According to the American Association of Intellectual and Developmental Disabilities (AAIDD, 2010) adaptive behavior can be defined as the conceptual, practical and social skills that are learned and executed by individuals on a daily basis. Examples of adaptive behaviors include language and communication skills, interpersonal skills, self-esteem, social problem-solving abilities, rule following and personal care skills. The Vineland ABS/II produces standard scores on five domains: Communication, Socialization, Daily Living Skills and Motor Skills. According to the *Individuals with Disabilities Education Act* (IDEA, 2004) an adaptive behavior measure is a necessary and integral part of assessment and diagnosis of DSM-5 defined intellectual disability (ID), and is also highly recommended when assessing for general developmental delay. In addition to its role in identifying adaptive behavior deficits associated with ID, the Vineland ABS/II is also used for treatment planning for individuals with intellectual and developmental disabilities and it has been extensively used in research with populations of typically developing and delayed children (Gleason & Coster, 2012).

Previous literature has documented that children with ASD demonstrate variability in subdomain performance on the Vineland ABS, e.g., differences in receptive and expressive language, and in personal, domestic, and community daily living skills (Burack & Volkmar, 1992). Carpentieri and Morgan (1996) found that in a sample of children with ASD (average age = 8 years) and children with intellectual disability (average age = 9 years), children with ASD demonstrated a pattern of poorer Socialization and Communication skills than Daily Living Skills as measured by the Vineland ABS. Further, Fenton et al. (2003) examined scores on the Vineland ABS in a sample of children ages 21 to 108 months with ASD and moderate to severe developmental disabilities. Fenton et al. (2003) found that as the level of disability increased (measured by the gap between chronological age and developmental age), the typical pattern of performance previously documented in the literature for children with ASD was not found. Fenton et al.'s finding is supported by Liss et al. (2001) who found that different adaptive behavior profiles were present in children with an average age of 60 months and "low-functioning autism," which was defined by a DSM-III diagnosis of autistic disorder and a nonverbal IQ score < 80, compared to children with an average age of 58 months and "high-functioning autism," defined by a DSM-III diagnosis of autistic disorder and a nonverbal IQ ≥ 80.

Some research also has revealed discrepancies between Vineland-II subdomain scores and corresponding domain scores from other cognitive assessments such as the Bayley Scales of Infant Development (Scattone et al., 2011), and other adaptive behavior measures, such as the Adaptive Behavior Assessment System (ABAS-II; Harrison & Oakland, 2003) (Milne &

McDonald, 2015) in samples of children with developmental delays. However, the factors that influenced these differences, such as item level content differences or differences in psychometric properties remain under debate and further research is needed to better understand why performance on similar constructs varies across assessments in children with developmental disabilities.

1.2.5 Domain-Specific Measures

In addition to general cognitive and adaptive functioning deficits, children with developmental disabilities often experience severe deficits in language and communication (Waterhouse & Fein, 1982). Research suggests that children who demonstrate spoken language delays early in life are at an increased risk of experiencing challenges throughout development, and that limited spoken language capabilities severely inhibit children's ability to interact with the world around them (Dale et al., 2003; Romski et al., 2002). Accordingly, language and communication are especially important domains to assess in this population. Several measures have been designed and utilized specifically to assess the domains of language and communication in young children.

1.2.5.1 The MacArthur-Bates Communicative and Developmental Inventory

The MacArthur-Bates Communicative and Developmental Inventory (MCDI) is the most commonly used caregiver report of communication skills in typically developing infants and toddlers 8-30 months of age, and it has also demonstrated reliability in populations with developmental disabilities (Plesa Skwerer et al., 2016). The MCDI measures early gesture use, vocabulary comprehension and vocabulary production and has been used pervasively in research and clinical contexts to measure early language abilities in very young children and toddlers with developmental disabilities (Bruckner et al., 2007; Dale, 1991; Miller, Sedey, & Miolo, 1995;

O'Toole & Fletcher, 2010). Importantly, the MCDI measures both vocabulary production and comprehension. Prior research indicates that in children under 2 years of age, a combination of ability in vocabulary production and comprehension is a stronger predictor of future language skills than vocabulary production alone (Fenson, 2007).

1.2.5.2 The Sequenced Inventory of Communication Development

The SICD is an examiner-administered assessment of language for children 4 to 48 months of age. This assessment is used to evaluate receptive and expressive language ability in children who demonstrate typical and delayed development. A unique and important characteristic of the SICD is that it includes a combination of items that are directly administered by a clinician and parent report items. This design can generate a more accurate picture of a child's performance, particularly in children that are more "difficult-to-test," such as children with autism or other developmental disorders (Hedrick, Prather, & Tobin, 2002). Additionally, the SICD utilizes visual and verbal stimuli that require responses in a variety of modalities including verbal, motor and visual; this design reduces penalization of children with extreme deficits in one modality. The SICD also includes items that are designed to assess typically developing children as young as 4 months, which permits most children, even those with very little ability, to achieve success early on during testing. As a result of these qualities, the SICD has been frequently utilized in research to measure language ability in young children and toddlers ages 12 months to 6 years with a range of developmental disabilities including general developmental delay and autism spectrum disorder (Brady, Marquis, Fleming, & McLean, 2004; Burchinal, Roberts, Hooper, & Zeisel, 2000; Calandrella & Wilcox, 2000; Lord & Pickles, 1996).

Some research has examined performance profiles on the SICD in populations with developmental disabilities. For example, Weismer et al. (2010) administered the SICD to a large sample of toddlers, ages 24 to 36 months with ASD, PDD-NOS and developmental delay (DD). Weismer et al. (2010) found that the group of children with DD (N=69) and PDD-NOS (N=78) demonstrated receptive language age equivalent scores that were higher than their expressive language age equivalent scores. In support of this finding, Calandrella and Wilcox (2000) found that the mean SICD receptive age equivalent in a sample of toddlers, ages 17 to 38 months with global developmental delays was slightly higher than the mean expressive age equivalent, however this was not assessed statistically. More research is needed to elucidate the language performance patterns of infants and toddlers with developmental delays across multiple parent informant and direct observation measures.

1.3 Research Questions and Hypotheses

The current study will examine performance within and across a series of direct observation and parent report developmental assessment measures that cover a range of abilities including expressive and receptive language, adaptive behavior, fine and gross motor skills, and visual reception in a sample of toddlers with a general developmental delay and less than 10 spoken words. We asked two broad questions pertaining to performance on developmental assessment measures in our sample of toddlers:

1. What are individual patterns of cognitive, adaptive and communication performance within a sample of toddlers with a significant developmental delay and less than 10 spoken words?

A. Are outliers present across assessment domains in the current sample? This exploratory analysis will contribute to the characterization of individual participants' assessment performances by identifying significant individual differences in performance relative to the average performance of other participants. We hypothesize that, in line with expectations according to a normal distribution, most participants will perform within the normal range with only a small number (approximately 5%) being outliers.

B. What are individual children's relative strengths and weaknesses across assessment domains, i.e. visual reception, fine and gross motor, receptive language, expressive language, daily living skills and socialization? We hypothesize that participants will demonstrate some relative strengths and weaknesses across developmental domains. This hypothesis is based on research indicating different patterns of performance between developmental domains in young children with developmental disabilities (Caselli et al., 1998; Fidler et al., 2007; Luyster et al., 2011; Singer Harris et al., 1997; Weismer et al., 2010).

2. What are group patterns of performance across assessment domains and measurement types?

A. How does average group performance compare across measures of cognitive, language and adaptive functioning in this sample of toddlers? For example, are Receptive and Expressive Language scores on the MSEL, the Vineland ABS/II, the SICD and the MCDI significantly related? We hypothesize that many domains will be significantly and moderately correlated, but that correlations may not be consistent across all measures and domains. We hypothesize this because of inconsistencies in previous research with toddlers with developmental disabilities that demonstrates some agreement across measures (Dale, 1991; Weismer et al., 2010), and some differences in domain performance across measures (Magiati & Howlin, 2001; Scattone et al., 2011).

B. How does performance compare across *types* of measures? Specifically, how does performance on certain domains from parent report measures compare to performance on direct assessment measures? We hypothesize that there will be agreement between parent report and direct observation assessment, but the agreement will be greater for measures of expressive language and motor skills than for measures of receptive language. This hypothesis is based on research indicating strong agreement between parent report and direct observation measures in young children with developmental disabilities (Björn et al., 2014; Dale, 1991; Scattone et al., 2011; Weismer et al., 2010), that appears to be moderated by development domain (Charman, 2004; Lyytinen et al., 1999).

2 METHODS

2.1 Participants

The current study included a sample of 129 children, mean chronological age = 29.77 months, *SD* = 5.04 months, range from 21 to 48 months, with a general developmental delay and severe spoken language impairment, operationally defined as an Expressive Language age equivalent score on the Mullen of less than 12 months and observed functional use of no more than 10 spoken words. Children also demonstrated gross motor skills that permitted the manipulation of a speech-generating device (SGD), and at least some beginning intentional communication ability, i.e. primitive vocalizations and gestures that refer to or request objects/events in their environments. All children's primary language was English. Medical diagnoses included seizures, genetic conditions and cerebral palsy. Children with primary etiologies of autism, hearing impairment/deafness or a speech and language delay without general developmental delay were excluded from the study. All children were assessed as part of their inclusion in one of two sequential longitudinal studies investigating the effectiveness of parent-implemented early language interventions (Romski et al., 2010; Romski et al., *manuscript in preparation*). This assessment took place prior to their enrollment in any intervention group. Participants for both studies were recruited within the metropolitan Atlanta area. Participants were recruited from various local professionals that had experience working with children with severe disabilities, including developmental pediatricians, pediatric neurologists, clinical psychologists and speech-language pathologists. The current sample consisted of 40 children of African American background (31%), 74 children of Caucasian background (57%), 10 children of Asian background (8%) and 5 children with a Multi-racial background (4%). Four parents identified their children as Hispanic (3%), 124 children were identified as Non-Hispanic (96%), and one child's parent did not report the child's ethnicity. The sample contained more male (N = 90, 70%) than female participants (N = 39, 30%).

For both longitudinal studies one parent for each child was selected to participate according to who could commit to participate in all of the study sessions. Parents who were selected to participate (N = 129) included 11 fathers and 118 mothers. Parents' mean age = 36.03 years, *SD* = 5.37 years, range from 21 to 45 years; 4 of the parents did not report their age, thus the mean age was calculated from the remaining 125 parents. This sample included 37 parents of African American background (29%), 74 parents of Caucasian background (57%), 9 parents of Asian background (7%) and the remaining 9 parents did not report their race. Additionally, 118 parents identified as Non-Hispanic (92%), 3 parents identified as Hispanic (2%), and 8 parents did not report their ethnicity. Parents' education levels varied, 1 parent did not attend high school (0.8%), 9 parents graduated from high school but did not attend college (7%), 18 parents completed at least some college (14%), 59 parents had bachelor's degrees (46%), 39 parents had graduate or professional degrees (30%) and 3 parents did not report their education background.

2.2 Procedures

A trained clinician administered a developmental assessment battery to each parent-child pair that included parent report and direct observation measures of general developmental level, adaptive behavior, communication, and visual-spatial and motor skills. All evaluations took place in the Toddler Language Intervention Project Lab at Georgia State University in a room with a child sized table and chair. During testing, one clinician was present in the room throughout each session. For children who would separate without distress, their parent was not present in the room during testing but was able to observe via a one-way viewing window from an adjacent room. For children who would not separate, their parent remained in the room during testing. Each testing session lasted approximately one hour and assessment batteries were completed over an average of five to six sessions. Testing was stopped if a child was not engaged, e.g. was crying or screaming during testing. Children were also given breaks throughout testing. All assessment sessions were videotaped.

2.3 Measures

Not all measures that were administered will be examined for the purposes of this study. The measures that will be utilized include the Mullen Scales of Early Learning (Mullen; Mullen, 1995), the Vineland Adaptive Behavior Scales (Vineland-ABS; Sparrow, Balla, Cicchetti, & Doll, 1984) and the Vineland Adaptive Behavior Scales, Second Edition (Vineland-II; Sparrow, Cicchetti, & Balla, 2005), the MacArthur-Bates Communicative Development Inventory (MCDI; Fenson, 2007) and the Sequenced Inventory of Communication Development (SICD; Hendrick, Prather, & Tobin, 1984). All of these measures have demonstrated adequate reliability and validity.

2.3.1 Mullen Scales of Early Learning

The Mullen is a clinician-administered developmental measure that assesses abilities across five domains: Gross Motor, Fine Motor, Visual Reception, Expressive Language and Receptive Language. Items assessing each of these domains are rated as either 0 or 1 depending on whether the child is able to complete the required tasks, such as naming objects pictured in a book, stacking blocks or matching pairs of objects. Raw scores from each domain are used to calculate T-scores (M = 50, *SD* = 10) and age equivalent scores in months. The Mullen also produces an overall cognitive measure, the Early Learning Composite (ELC) that theoretically equates to a measure of “g,” underlying performance across all domains. Mullen (1995) reported median Cronbach’s alpha coefficients for Mullen subtests between .75 and .85, test-retest correlation coefficients between .76 and .83, interscorer correlation coefficients between .91 and .99 and strong construct and concurrent validity.

While the original norming sample for the Mullen did not include children with developmental disabilities, a recent psychometric investigation by Swineford et al. (2015) revealed that the Mullen demonstrated strong construct validity in a sample of children with ASD and non-spectrum developmental delays. More specifically, a confirmatory factor analysis revealed that each of the Mullen subscales mapped onto a latent "g" factor of general cognitive level, approximated by the ELC. When the sample was divided by diagnosis, CFA loadings onto "g" were significantly less (but to a small degree) in children with ASD than children with nonspectrum delays. However, the factor loadings in the ASD group were still strong (.69-.83), and similar to those found in other cognitive assessments. Swineford et al. (2015) also found that the Mullen demonstrated strong convergent and divergent validity in the ASD and nonspectrum delays groups.

2.3.2 Vineland Adaptive Behavior Scales

The Vineland is a parent interview that measures adaptive functioning across four domains: Communication, Daily Living Skills, Socialization and Motor Skills. The Vineland has two versions, one that can be completed by parents or caregivers independently, and one that is administered to parents or caregivers via a trained clinician. The Vineland survey form, which is administered by a clinician, was used in this study. All items are scored on a 3-point scale of behavior frequency, with 0 indicating *never*, 1 indicating *sometimes or partially*, and 2 indicating *usually*. In 2005 Sparrow, Cicchetti, and Balla developed a new version of the Vineland ABS, the Vineland-II. In this study parent-child dyads participated in one of two longitudinal studies; the first of those studies began prior to the development of the Vineland-II and therefore those parents were administered the original Vineland ABS. The second study began after 2005, so parents who participated in that study were administered the Vineland-II. Changes to the updated version of the Vineland were made to improve measurement in very young children and older adults. The primary modifications made to the original version included additional items to address the following areas: ability to start and maintain conversations and spoken language skills in the Expressive and Receptive subdomains, independent living skills in the Daily Living Skills domain, and use and comprehension of nonverbal communication during social interactions in the Socialization domain. In the Vineland-II manual, large correlations were reported between domain standard scores for a sample of 24 children birth through age 2, and 29 children 3 through 6 years. For the sample of children birth through age 2, correlations for the Communication, Daily Living Skills, Socialization, Motor Skills and Adaptive Behavior Composite domains were r = .65, .75, .85, .91 and .82, respectively. For the sample of children ages 3 to 6 years, correlations for the Communication, Daily Living Skills, Socialization, Motor Skills and Adaptive Behavior Composite domains were r = .85, .91, .94, .86 and .91, respectively. Mean differences between domain scores across the versions were small, except for the Daily Living Skills domain (Vineland ABS SS mean = 78.7, Vineland-II SS mean = 87.8). In general, mean domain scores from the Vineland-II were slightly higher than those from the Vineland ABS. All analyses in this study that incorporate Vineland domain scores (Communication, Daily Living Skills, Socialization, Motor Skills and Adaptive Behavior Composite) combine data from children whose parents completed the Vineland ABS and Vineland-II versions in the same analysis.

The Vineland ABS demonstrated good reliability and validity (Sparrow et al., 1984). Specifically, Sparrow et al. reported split-half reliability coefficients ranging from .83 to .94

across all domains and slightly lower reliability coefficients for subdomains ranging from .69 to .84, test-retest reliability coefficients for children from birth through 4 years, 11 months ranging from .78 to .92, and interrrater reliability coefficients ranging from .62 (for the Socialization domain) to .78 for the entire normative sample. Sparrow et al. also reported strong validity illustrated by moderate correlations between the Vineland ABS and other adaptive behavior measures, with higher correlations between these measures than measures of cognitive functioning, such as the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). Factor analysis also confirmed the appropriate organization of all of the Vineland ABS's subdomains within their respective domains (i.e. Communication, Daily Living Skills, Socialization and Motor Skills) for children ages 2 to 3 years old. Similarly, Floyd et al. (2015) reported Cronbach's alpha coefficients above .80 for the Vineland-II, test-retest correlation coefficients greater than .90 and adequate interrater reliability and validity. Contrastingly, Plesa Skwerer et al. (2016) found that in a sample of children with ASD, Vineland-II Receptive Language subdomain scores were not correlated with other language measures (i.e. the PPVT-4, a parent checklist of vocabulary words and the Raven matrices). However, the researchers hypothesized that this finding could be attributed to the Vineland's inclusion of questions that measure other aspects of language including pragmatics, i.e. it is not a measure of single-word vocabulary comprehension like the other language measures used in the study. This is an important finding to be cognoscente of when interpreting performance on this measure, and when comparing language performance between the Vineland ABS/II and other language measures.

2.3.3 MacArthur-Bates Communicative Development Inventory

The MacArthur-Bates Communicative Developmental Inventory (CDI) is a parent report measure that is used to assess early receptive and expressive communication abilities in toddlers 8 to 30 months of age. In this study the MacArthur-Bates CDI: Words and Gestures form, developed for children ages 8 to 18 months of age, was administered to parents. This form consists of two parts: Part I, which assesses early signs of comprehension and includes a 396-item checklist of vocabulary words that parents indicate their child either “understands” or “understands and says,” and Part II, which assesses the use of symbolic communicative gestures and actions such as playing peek-a-boo, waving good-bye, putting on a hat and imitating adults’ actions. The current study will utilize the MCDI: Words and Gestures Words Understood and Words Produced categories as independent variables in data analyses. Fenson et al. (2007) reported Chronbach’s alpha coefficients of .95 and .96 for Words Understood and Words Produced, respectively. In regards to predictive validity, the MCDI: Words and Gestures form was administered to a group of parents with toddlers ages 10-16 months and then again approximately 6 months later. Words and Gestures at Time 1 and Time 2 demonstrated strong significant correlations, $r \geq .65$ for all age groups except the 12 month age group where $r = .38$. The MCDI has also demonstrated strong concurrent validity with other language assessments (Nordahl-Hansen et al., 2014) including strong correlations with clinician-administered direct assessments of receptive and expressive language skills, such as the Bayley Scales version-III (Bayley & Reuner, 2006) in young children 12 to 18 months of age (Björn et al., 2014) and laboratory observations of vocabulary use (Miller et al., 1995; Thal, O’Hanlon, Clemmons, & Fralin, 1999).

2.3.4 Sequenced Inventory of Communication Development

The Sequenced Inventory of Communication Development (SICD; Hedrick, Prather, & Tobin, 1984) is an assessment tool used for evaluating language and communication skills in young children who are typically developing and in those with developmental delays. The SICD includes both clinician-administered direct testing, and parent report items. The SICD is composed of two sections: Receptive language, which measures a child's ability to decipher, recognize and comprehend speech sounds and words, and Expressive language, which assesses a child's ability to produce and appropriately use verbal output to communicate. The SICD has demonstrated strong reliability and validity (Hedrick et al., 1984). For example Hedrick et al. reported that sufficient interrater reliability was achieved on a subset of the normative sample; the average agreement between raters for whether individual items should be rated as "pass" or "fail" was 96% (range = 90% to 100%). Test-retest reliability also exhibited sufficient agreement; the average percentage agreement between administrations, which were approximately one week apart, was 93% (range = 88 to 99%). Validity of the measure was also reported; correlation coefficients between the SICD and other measures of language (e.g. the PPVT) ranged from .75 to .80. These coefficients demonstrated a high enough correlation to support that the construct of language was being measured, yet low enough to also support that the SICD makes a unique contribution to the measurement of language and is not identical to the other language measures.

3 RESULTS

3.1 Descriptive statistics and patterns of outliers across measures

All participants (N = 129) were administered a developmental battery that measured a broad range of abilities including motor, language and communication, and adaptive functioning skills. Average performance fell at least two standard deviations below the mean across all five developmental domains on the Mullen Scales of Early Learning, i.e. Gross Motor, Fine Motor, Visual Reception, Expressive Language and Receptive Language, and between one and a half and two standard deviations below the mean across domains of adaptive functioning on the Vineland ABS/II, i.e. Communication, Daily Living Skills, Socialization and Motor Skills. Additional descriptive statistics are reported in Table 1. To address research question 1A, the data were inspected for outliers. Boxplots were created to investigate potential ($z > 1.96$), probable ($z > 2.58$) and extreme ($z > 3.29$) outliers on all measures. A total of 49 children demonstrated an outlier on at least one assessment subdomain. The majority of children demonstrated only one outlier (N=26, 53%), with fewer demonstrating two outliers (N=16, 33%), and only seven children (14%) demonstrated outliers across three or more test subdomains. Boxplots of Mullen standard scores revealed that the gross motor domain contains five (3.9%) potential outliers and one (0.8%) extreme outlier, the fine motor domain contains seven (5.4%) potential, one (0.8%) probable and one (0.8%) extreme outlier, the visual reception domain contains four (3.1%) potential and two (1.6%) extreme outliers, the expressive language domain contains four (3.1%) potential, three (2.4%) probable and two (1.6%) extreme outliers, and the receptive language domain contains four (3.1%) potential and four (3.1%) probable outliers. On a parent report measure of expressive and receptive language (MCDI), participants were reported to have an average of 8.6 spoken words, SD = 12.8, and to understand an average of 118.7 words, SD = 97.7; there was a notably large variation in the parent-reported number of words children are able to understand. Boxplots of MCDI raw scores revealed that the expressive subscale (words spoken) contains two (1.6%) potential, two (1.6%) probable and three (2.3%) extreme outliers, and the receptive subscale (words understood) contains four (3.1%) potential and three (2.3%) probable outliers. According to a parent report and clinician-administered language assessment (SICD), participants had an average expressive language age equivalent of 12.6 months, SD = 4.9 months, and an average receptive language age equivalent of 17.1 months, SD = 6.6 months. Boxplots of SICD age equivalents revealed that the expressive subscale contains five (3.9%) potential outliers, and the receptive subscale contains four (3.1%) potential, two (1.6%) probable and one (0.8%) extreme outlier. On a parent report measure of adaptive functioning ability, children performed at least one to two standard deviations below the mean across composite domains for communication, daily living skills (DLS), socialization skills and motor skills. Boxplots revealed one (0.8%) potential and one (0.8%) probable outlier on the Communication domain, four (3.1%) potential, two (1.6%) probable and one (0.8%) extreme outlier on the DLS domain, six (4.7%) potential outliers on the Socialization domain, and two (1.6%) probable and one (0.8%) extreme outlier on the Motor Skills domain.

Table 3.1 Participant performance across assessments

	N	Min	Max	M	SD
Mullen					
Gross motor[1]	108	19	66	29.1	11.0
Visual reception	129	19	73	30.3	12.4
Fine motor	129	19	54	24.9	8.2
Receptive	129	19	59	26.8	11.1
Expressive	129	19	39	21.3	4.0
ELC	129	48	97	57.6	11.2
SICD					
Receptive	128	4	40	17.0	6.6
Expressive	129	3	24	12.6	4.9
MCDI					
Receptive	125	2	395	118.7	97.7
Expressive	125	0	75	8.6	12.7
Vineland ABS					
Communication SS	74	55	82	67.1	6.5
Daily Living SS	74	48	92	69.5	8.5
Socialization SS	74	55	92	72.7	8.9
Motor skills SS	74	30	111	72.5	13.0
ABC	74	44	88	65.0	7.3
Vineland II					
Communication SS	55	6	86	67.5	12.8
Daily Living SS	55	29	96	74.7	12.8
Socialization SS	55	61	93	75.9	7.9
Motor skills SS	55	52	96	77.9	10.2
ABC	55	51	87	70.9	7.5

Note. ELC =Early Learning Composite (SS); ABC=Adaptive Behavior Composite (SS). T-scores are reported for each Mullen subscale (M=50, SD = 10). Age equivalents are reported for both SICD subscales. Raw scores are reported for MCDI Receptive = number of words understood, and MCDI Expressive = number of spoken words.

[1]20 children do not have MSEL gross motor SS's because they were 34 months of age or older and normative data for this subscale is not available for children over 33 months of age, 1 child's MSEL did not include this subscale because it was administered by an outside psychologist.

3.2 Individual strengths and weaknesses across domains

To participate in the original study, all toddlers were required to demonstrate marked delays in expressive language, i.e. less than 10 spoken words. However, their skills in other developmental domains were unrestricted at study entry, which raises the question of how these children performed on measures of other abilities. To address research question 1B, the standard error of differences (SE_{diff}) was calculated to evaluate significant individual patterns of strengths and weaknesses across developmental domains on the Mullen Receptive, Fine Motor, Gross Motor and Visual Reception domains in comparison to the Mullen Expressive Language domain. According to guidelines outlined by Coaley (2014), SE_{diff} scores were calculated using the following equation: square root $[(SE_{m1})^2 + (SE_{m2})^2]$, where SE_{m1} and SE_{m2} are the standard errors of measurement for score 1 (e.g., Mullen Expressive Language) and score 2 (e.g., Mullen Receptive Language). Differences between individual participants' expressive language T-scores and T-scores on other Mullen domains that are equal to two and three SE_{diff} units are reported. Results revealed that 25 individuals demonstrated a difference between their expressive language and fine motor T-scores equal to two SE_{diff} units, with 21 demonstrating differences equal to three SE_{diff} units, 34 individuals demonstrated a difference between their expressive and receptive language T-scores equal to two SE_{diff} units, with 18 demonstrating differences equal to three SE_{diff} units, 38 individuals had two SE_{diff} units between their expressive language and gross motor scores, with 25 of those being three SE_{diff} units apart, and 48 individuals exhibited two SE_{diff} units between their expressive language and visual reception T-scores, with 31 being three SE_{diff} units apart. All of these differences reflected relative weaknesses in expressive language compared to the other domains, except three individuals who demonstrated stronger expressive language than gross motor ability (two with a difference of two SE_{diff} units, one with a difference of three SE_{diff} units). Notably, these individuals all had a diagnosis of Down syndrome.

3.3 Relations between assessment measure domains and subdomains

To address research question 2A, scatterplots and Pearson's r correlations were conducted to investigate the relations between performance on developmental, language and adaptive functioning measures. A scatterplot of overall developmental level, measured by the Mullen Early Learning Composite (ELC), and overall adaptive level, measured by the Vineland ABS/II Adaptive Behavior Composite (ABC) revealed a roughly positive linear relation, see Figure 1. A simple linear regression revealed that performance on the Mullen ELC was significantly linearly related to performance on the Vineland ABS/II ABC, $R^2 = .23$, $t = 14.86$, $p < .001$. However, visual inspection and a local regression estimation line (LOESS) of the scatterplot also revealed a potentially curvilinear relation between the Mullen ELC and Vineland ABS/II ABC. Therefore, a hierarchical multiple regression was conducted to assess the relative contribution of adding a quadratic regression line to the data. The regression revealed that while a significant linear relation is present between Mullen ELC and Vineland ABS/II ABC scores, a significant amount of additional variance is explained by adding a quadratic parameter to the regression equation, $\Delta R^2 = .06$, $p = .001$. It is important to take note of the large variance of performance on the Vineland ABS/II ABC at differing overall developmental levels, particularly for children who performed 3 to 4 standard deviations below the mean on the Mullen ELC (ELC SS $\leq$ 55). Children who performed at the lower end on the Mullen ELC (N = 70) demonstrated a wide range of parent-reported adaptive skills, i.e. Vineland ABS/II ABC standard scores ranging from 44 to 83. There was also large scatter for those children who performed within only one standard deviation of the mean on the Mullen (Mullen ELC $\geq$ 85; N = 7), with Vineland ABS/II ABC scores ranging from 60 to 88.

Pearson's r correlations were conducted for all domains on parent report and clinician-administered assessment measures. Performance on developmental measures, language measures and adaptive functioning measures were highly and positively correlated across most domains between and within tests, with the greatest exception to this being performances on motor domains. There were two small, non-significant negative correlations between performance on the Vineland ABS/II Motor Skills domain and the SICD and Mullen Receptive Language domains. The Vineland ABS/II Motor Skills domain was also not significantly correlated with the MCDI Receptive Language domain, or the Expressive Language domains of the Mullen or the SICD. The Mullen Gross Motor domain was not significantly related to the MCDI or Mullen Receptive Language domains, or to Vineland ABS/II Socialization or Communication domains, and Mullen Fine Motor performance was not significantly related to Vineland ABS/II Communication performance. Interestingly, while the Expressive Language domain of the MCDI was significantly correlated with all domains on the Mullen and the SICD, it was not significantly correlated with the any areas of adaptive functioning on the Vineland ABS/II. To further investigate the small relation between parent-reported language and communication functioning (measured by the Vineland ABS/II Communication SS), and parent-reported number of spoken words (measured by the MCDI expressive raw score), MCDI Expressive Language (EL) scores were correlated with Vineland ABS/II Receptive Language (RL) and Expressive Language (EL) raw scores. Because different versions of the Vineland were used across the two studies included in the current analyses, and thus raw scores were not equally scaled across all participants, correlations were run separately for participants recruited from each study. This analysis revealed large, significant correlations between MCDI EL and Vineland-II RL and EL, *r* = .48, *p* >.001, and *r* = .61, *p* >.001, respectively for participants from the second study (N=55). However, correlational analysis for participants in the first study (N = 74) revealed a small, significant relation between MCDI EL and Vineland ABS RL, *r* = .32, *p* = .007, and a small, non-significant correlation between MCDI EL and Vineland ABS EL, *r* = .22, *p* = .07. All correlation results are reported in Table 2.

Table 3.2 Correlations across assessment measures

	Mullen					SICD	
	GM	VR	FM	RL	EL	RL	EL
Mullen GM	1.00						
Mullen VR	0.27**	1.00					
Mullen FM	0.66**	0.57**	1.00				
Mullen RL	0.13	0.65**	0.38**	1.00			
Mullen EL	0.27**	0.23**	0.36**	0.37**	1.00		
SICD RL	0.11	0.61**	0.37**	0.87**	0.38**	1.00	
SICD EL	0.31**	0.30**	0.36**	0.44**	0.68**	0.44**	1.00
MCDI RL	0.14	0.43**	0.32**	0.62**	0.28**	0.64**	0.36**
MCDI EL	0.21*	0.18*	0.29**	0.30**	0.55**	0.28**	0.45**
VABS Com	-0.06	0.27**	0.08	0.46**	0.22*	0.41**	0.36**
VABS DLS	0.32**	0.29**	0.35**	0.21*	0.24**	0.20*	0.29**
VABS Soc	0.07	0.40**	0.18*	0.40**	0.25**	0.39**	0.32**
VABS MS	0.66**	0.18*	0.56**	-0.03	0.15	-0.02	0.17

Note. Raw scores were used for the Mullen; age equivalents were used for the SICD. GM= Gross Motor, VR=Visual Reception, FM=Fine Motor, RL=Receptive Language, EL=Expressive Language. ** Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).

Table 3.3 Correlations across assessment measures cont'd

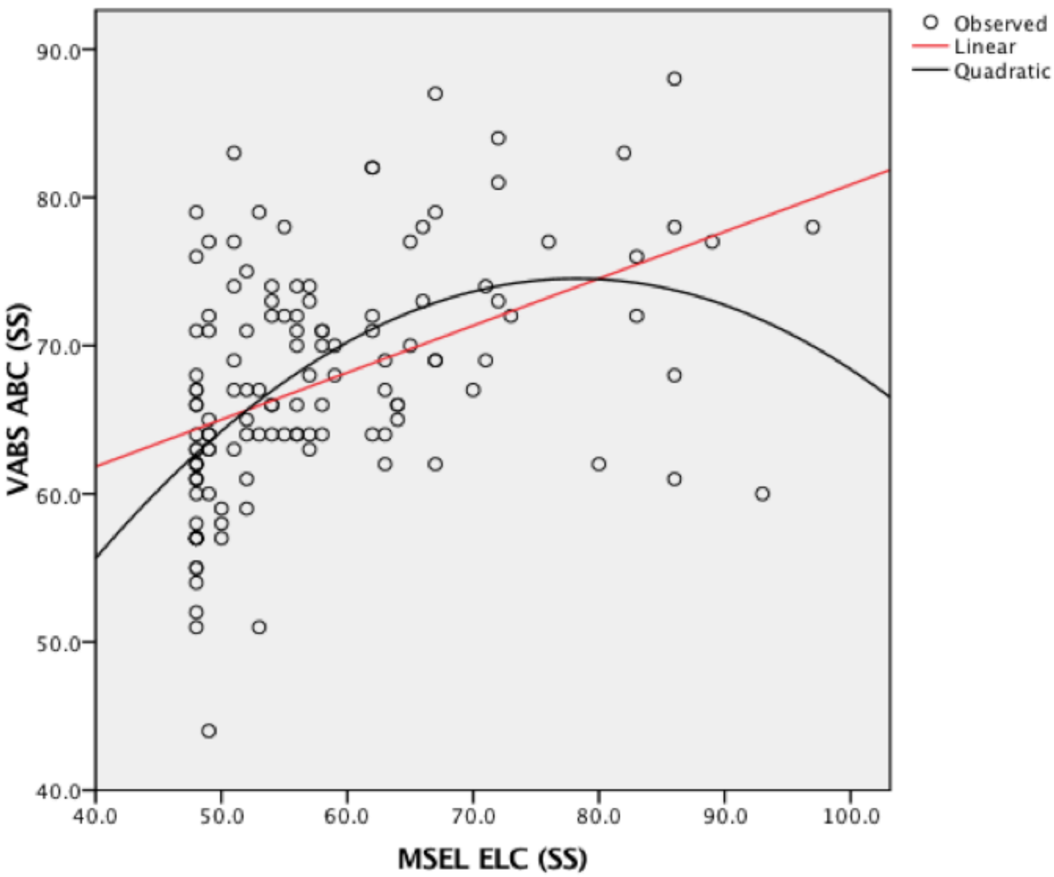
	MCDI		Vineland ABS/II			
	RL	EL	Comm	DLS	Soc	MS
MCDI RL	1.00					
MCDI EL	0.44**	1.00				
VABS Com	0.32**	0.14	1.00			
VABS DLS	0.19*	0.16	0.50**	1.00		
VABS Soc	0.24**	0.09	0.65**	0.57**	1.00	
VABS MS	0.03	0.13	0.23**	0.52**	0.30**	1.00

Note. Raw scores were used for the MCDI; standard scores were used for the Vineland ABS/II domains. RL=Receptive Language, EL=Expressive Language, Comm=Communication, DLS=Daily Living Skills, Soc=Socialization, MS=Motor Skills. * Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

3.4 Relations between parent report and clinician-administered measures

To address research question 2B, paired samples t-tests were conducted to evaluate the relations between children's performances on a clinician-administered assessment (Mullen) and a parent report interview (Vineland ABS/II) across four developmental subdomains: gross motor, fine motor, receptive language and expressive language. Because these subdomains are not comparable across the two Vineland versions, t-tests were run separately for each study (study 1, N= 74; study 2, N = 55). To run t-tests using these measures, it was necessary to create standardized scores that could be appropriately compared across both assessments. Therefore, z-scores were calculated for participants' scores on each domain and measure using the following equation: (individual score – sample mean)/sample standard deviation. For both studies, there were no significant differences between parent-reported and clinician-assessed abilities, except for within the gross motor domain, $t = -2.99$, $p = .004$ and $t = 3.72$, $p < .001$ for study 1 and study 2, respectively. For study 1, the average z-score for clinician-administered (Mullen) gross motor performance was lower than the average z-score for parent-reported (Vineland) gross motor skill. For study 2, the reverse was true, with average clinician-administered gross motor functioning being higher than parent-reported gross motor skill. However, according to Cohen (1988, 1992), effect sizes were large for the correlations between parent report and clinician-evaluated gross motor skills in both studies. This suggests that while average z-scores were different between the assessment types, parent and clinician ratings were highly related across participants. Large, significant relations were also found on measures of fine motor, receptive language and expressive language domains for study 1, $r = .72$, .74, and .42, and study 2, $r = .69$, .77, and .57, respectively. However, the effect size for the expressive language domain was noticeably smaller than the effect sizes for the receptive language and motor domains. To further investigate the weaker relation between parent-reported and clinician-administered expressive language abilities, the sample was divided into two groups: 1) children whose performance fell above the median Mullen Expressive Language raw score and 2) children whose performance fell below the median raw score. For study 1, receptive language abilities (as measured by Mullen Receptive Language standard scores) of children in Group 1 (N = 44), M = 29.14, SD = 11.6, range = 19-59, were higher than children in Group 2 (N = 30), M = 25.20, SD = 13.04, range = 19-59. Paired samples t-tests comparing parent-reported and clinician-evaluated expressive language abilities were conducted on these groups separately. T-tests revealed that there was a significant difference between parent report and clinician-administered assessment of expressive language in both groups; children who performed below the median, $t = -2.01$, $p = .05$, and children who performed above the median, $t = 4.91$, $p < .001$. For children who performed above the median (Group 1), clinicians reported greater expressive language than parents; for children who performed below the median (Group 2), parents reported greater expressive language abilities than clinicians. Interestingly, the relation between parent report and clinician assessment was non-significant and weaker for children who performed below the median, $r = .24$, $p = .212$

than for children who performed above the median, *r* = .40, *p* = .007. For study 2, the average Mullen Receptive Language (standard score) for children in Group 1 (N = 27), M = 28.14, SD = 10.06, range = 19-54, was higher than for children in Group 2 (N = 28), M = 23.36, SD = 13.04, range = 19-45. Paired samples t-tests comparing parent-reported and clinician-evaluated expressive language abilities were conducted on these groups separately. T-tests revealed that there was a significant difference between parent-reported and clinician-administered assessment of expressive language in Group 2, *t* = -2.74, *p* = .011, but not for children in Group 1, *t* = 0.89, *p* = .383. Similarly to study 1, children in study 2 who performed above the median (Group 1), demonstrated greater clinician-assessed than parent-reported expressive language; children who performed below the median (Group 2), demonstrated greater parent-reported than clinician-administered expressive language abilities. The relation between parent report and clinician assessment measures was also non-significant and weaker for children who performed below the median, *r* = .18, *p* = .371 than for children who performed above the median, *r* = .34, *p* = .081. Overall, it appears that for children with lower expressive language performance on the Mullen, there is a weaker relation between their testing performance and parent-reported expressive language abilities.



Note. The scores plotted above are standard scores (M=100, SD=15). ELC = Early Learning Composite, ABC = Adaptive Behavior Composite. Linear regression line estimation is illustrated in red, F = 37.0, p < .001 quadratic regression estimation is illustrated in black, F = 25.56, p < .001.

Figure 3.1 Scatterplot of Mullen ELC and Vineland ABS/II

4 DISCUSSION

This study addressed profiles of abilities in a sample of toddlers with less than 10 spoken words and developmental disabilities across multiple domains of functioning and types of assessment methods. Regarding individual performances, about one third of the sample exhibited outliers on at least one domain (e.g., motor skills, language, adaptive functioning). Over half demonstrated an outlier in a single domain exclusively, with only 14% demonstrating outliers in three or more domains. This finding suggests that the majority of these children were not just the most or least impaired overall, but had specific areas of significant strength or weakness relative to the average performance of other young children with limited spoken language. Further analysis of individual performance profiles indicated that 48% of the sample exhibited significant differences between expressive language ability and ability in another domain, with the majority demonstrating relative strengths in receptive language and fine motor skills compared to expressive language. Regarding group performance patterns, average performances across developmental domains, and across measurement types were highly related. However, relations between domains and measurement types varied according to degree of expressive language impairment, i.e. children with the lowest expressive language performances tended to demonstrate less correspondence between parent report and clinician-administered measures.

4.1 Individual performance patterns

Hypotheses for research questions 1A and 1B were supported. Inspection of the data revealed that performances on most developmental domains contained ≤ 5% outliers, in line with the original hypothesis that approximately 95% of the participants' performances would fall within a normal range. A few exceptions were the fine motor, receptive language and expressive language domains of the Mullen, which contained 6-7% outliers. Over half (64%) of the participants exhibited a significant difference between observed expressive language ability and performance in another domain, which supports our original hypothesis that participants would demonstrate relative strengths and weaknesses in their abilities across developmental domains. These findings underscore the presence of varying performances across different areas of functioning, even in a sample of toddlers with general developmental delay and similarly impaired spoken language. Other researchers have encouraged consideration of individual performance profiles in evaluating children with developmental disabilities. For example, Scattone et al. (2011) compared performance on the Vineland-II and Bayley-III in a sample of toddlers with developmental delays, and found that while mean standard scores and age-equivalents between parent report and direct testing were highly correlated, large within-subject discrepancies between the measures were present. Scattone and colleagues concluded that while clinician-administered and parent report measures appear to correspond overall, individual performances across domains can be inconsistent in ways that are potentially meaningful for characterizing ability level and for treatment planning. In our study, three toddlers demonstrated stronger expressive language than gross motor skills, with one demonstrating stronger expressive language measuring two SE_{diff} units. This finding may be best understood by considering this child's diagnosis of Down syndrome (DS). Some research suggests that early motor skill acquisition is delayed in individuals with DS (Palisano et al., 2001; Pereira, Basso, Lindquist, Silva, & Tudella, 2013). Regarding language, many individuals with DS have language difficulties that persist into adulthood, (Chapman & Hesketh, 2000; Fowler, 1990; Laws & Hall, 2014), and they tend to demonstrate stronger receptive than expressive language skills starting in early toddlerhood (Fidler et al., 2007; Miller & Miller, 1999). However, early language milestones, e.g. first single word acquisition, tends to be achieved at similar rates compared to mental-age matched peers without DS (Chapman, 1997; Fowler, 1990). The participants of interest in this study had significantly weaker gross motor functioning than expressive language, and although expressive language was still delayed, it was on par with their nonverbal mental ages. This pattern is in line with previous research on toddlers with DS, who exhibit early expressive language that is typical for overall cognitive level, with specific language weaknesses becoming more apparent later in development. Still, these children represent a unique profile that was different from other toddlers in the study who also had a diagnosis of DS (N = 26). It is important for researchers and clinicians to consider unique profiles of strengths and weaknesses when evaluating and treating young children with developmental disabilities, even those with a shared genetic syndrome. This point has been highlighted by other researchers who emphasize that developmental profiles are "a quality of an individual, not a group of individuals (p. 124)," and have outlined methodological and conceptual issues with drawing conclusions about syndrome characteristics according to average group performance, and not individual abilities (Mervis & Robinson, 1999; Mervis, Robinson, Levy, & Schaeffer, 2003).

4.2 Group performance patterns

Hypotheses for research question 2A were supported. Initial whole group analysis of relations between overall developmental level and adaptive functioning suggested that performances in these domains are moderately related in a generally linear fashion. This is in line with previous research findings demonstrating moderate and significant correlations between performance on developmental, or cognitive, and adaptive functioning measures in children with developmental disabilities, e.g. ASD and ASD with "minimal" spoken language (Frost, Hong, & Lord, 2017; Ray-Subramanian, Huai, & Ellis Weismer, 2011). However, further investigation in this study indicated that the relation between these domains was better described by a curvilinear trend, revealing a more complex relation. At the lower end of overall developmental level (as measured by the Mullen ELC), there was a great deal of variability in parent-reported overall level of adaptive functioning. These findings could reflect true differences in adaptive functioning skills among children at similar developmental levels. However, differences this large in adaptive functioning wouldn't be expected for toddlers performing within a few standard scores of one another on a developmental measure. Previous research suggests a strong relation between level of neurodevelopmental impairment and aspects of functioning measured by the Vineland ABS such as language, socialization and behavioral difficulties (Ross & Weinberg, 2006). One study by Fidler, Hepburn, and Rogers (2006) demonstrated significant, large correlations between performance on the Mullen and Vineland ABS in a sample of toddlers with developmental delays. Thus, there may be alternative interpretations for the current findings. It is possible that the nonlinearity of the relation between overall developmental level and adaptive functioning was influenced by measurement challenges associated with evaluating toddlers with developmental disabilities. Parent report of adaptive functioning may be especially difficult to ascertain in a population of children with significant disabilities whose independent adaptive skills are limited. For example, parents may work more to meet their children's needs in order to accomplish daily tasks, and thus have more difficulty, or less opportunities to assess independent versus assisted abilities. Prior research supports some inaccuracies of parent report, e.g. a tendency for parents to over-report comprehension skills in children with developmental disabilities (Charman, 2004; Scattone et al., 2011). Conversely, clinician-administered assessments are scored according to information collected during relatively short periods of time in an unfamiliar testing environment with unfamiliar adults; this setting may be especially challenging for young children with developmental disabilities (Baker et al., 2002; Hauser-Cram

& Woodman, 2016). These aspects of direct testing may have contributed to underestimation of abilities for some of the children whose parents reported higher adaptive functioning at home and in more naturalistic contexts.

Regarding relations between specific developmental domains, findings from this study indicated strong, positive relations between most domains measured on cognitive, language and adaptive functioning assessments. One deviation to this pattern was the presence of small and non-significant correlations between parent-reported and clinician-administered motor and language skills. It is commonly accepted that early motor skills play an influential role in the development of language (Iverson, 2010). Several large scale review studies have also confirmed that motor difficulties often co-occur in children with language impairments (E L Hill, 2001; Rechetnikov & Maitra, 2009). However, as Hill (2010) pointed out, the relation between motor and language development is complex, and much remains to be understood about how these domains interact in children with developmental disabilities. The sample of toddlers in this study did not exhibit a clear relation between their motor and language abilities. Thus, the current findings support Hill's claim that the relation between these domains is not entirely straightforward, and may not be linear across all children with developmental disabilities.

Hypotheses for research question 2B were partially supported. Our findings demonstrated a strong overall correspondence between parent report and direct assessment measures. This is in line with previous research that has found agreement among parent-reported and clinician-administered measures in children with developmental disabilities, particularly with regard to measures of expressive and receptive language (Björn et al., 2014; Dale, 1991; Scattone et al., 2011; Weismer et al., 2010). We hypothesized that relations between parent-reported and clinician-administered abilities would be strongest for expressive language and gross motor domains, and weaker for receptive language. In line with our hypothesis, relations between measurement types for gross motor skills were the strongest, with fine motor ability following close behind. Although gross motor skills were highly correlated across measurement types in both studies, average parent-reported gross motor ability was significantly higher than clinician-assessed gross motor performance in study 1, and significantly lower than clinician evaluation of gross motor skills in study 2. Differences between study 1 and 2 may be related to the use of the Vineland ABS in study 1 versus the Vineland-II in study 2. The Vineland-II includes twice as many gross motor items as the Vineland ABS and is administered as a separate set of questions from the fine motor scale (these scales were administered simultaneously for the Vineland ABS). It is possible that these aspects differentially affected average parent report relative to average clinician administered gross motor skills. For example, the use of fewer items on the Vineland ABS may have reduced opportunities for parents in study 1 to describe their children's specific, independent gross motor skills and consequently propagated some over-reporting. Significant difference between measurement types in both studies may reflect differing reference points for parents and clinicians when evaluating gross motor skills. For example, when responding to interview questions, parents might compare their toddlers' gross motor skills to that of a higher functioning older sibling, and consequently under-report some abilities. Or, for parents with limited exposure to other children without developmental disabilities, they may over-report the typicality of their child's abilities. It is also possible that the clinicians administering the Mullen in this study, who are experienced in working with young toddlers with developmental disabilities, were able to use behavioral strategies and techniques to aid children in engaging in non-preferred gross motor tasks that they may be more resistant to demonstrating at home. Alternatively, this subscale was often performed towards the end of the developmental evaluation, and therefore factors such as fatigue and increased noncompliance could have contributed to under-performance on the gross motor domain. These are all important factors to consider when using individual performance on either parent report or clinician-administered measures to make diagnostic, service eligibility or treatment planning decisions.

In contrast to our original prediction, the relation between measurement types was weaker for expressive language than receptive language. To further investigate the relation between measurement types for expressive language, the sample was divided into two groups according to clinician-evaluated expressive language ability. This analysis revealed that parent report for the group of children with the lowest expressive language performance was significantly different from, and tended to be higher than clinicians' ratings. One possible explanation for this finding is that due to the already limited spoken language of this subgroup, the restricted time frame of formal standardized testing was insufficient to capture the participants' true expressive language abilities. Alternatively, it is possible that parents of children with little spoken language have more difficulty describing their children's expressive language level because examples of these skills occur at less frequent intervals than for children with more spoken language, making it more challenging to accurately report. A significant difference between parent-reported and clinician-administered expressive language was also demonstrated in the group of children with greater expressive language (for study 1 only), with average parent report being lower than clinician assessment in both studies. However, in study 1 the group of children with greater expressive language demonstrated a moderate-to-large, significant correlation between parent and clinician-administered measures, unlike the group of children with lower expressive language, where measurement types exhibited a non-significant, small-to-moderate relation. In study 2, parent-reported and clinician-administered measures exhibited non-significant correlations for both groups, but for the group of children with greater expressive language, the correlation was larger. Taken together, these findings underscore the importance of conducting multi-method, multi-informant assessments of language; especially for children with very limited or no spoken words, as information from only one source may under- or over-estimate a child's abilities and misinform treatment planning.

4.3 Limitations

There are several important limitations of this study that must be acknowledged. First, a comparison group, typically developing or with other developmental disabilities, was not included. Comparison groups may have been useful for drawing conclusions about the specificity of the developmental characteristics observed in the current sample of toddlers, and for quantifying the level of impairment in comparison to typically developing toddlers. However, designing appropriate comparison groups for individuals with developmental disabilities is challenging, especially for toddlers whose abilities are rapidly changing and in nonlinear ways during the first months of development (Mervis & Robinson, 1999; Mervis et al., 2003).

Second, because this study includes data from two time-separate projects, there were some differences versions of assessment measures used across projects, e.g. the Vineland ABS was used in the first study and the Vineland-II was used in the second. This resulted in some challenges for combining information across studies. One analysis revealed that performances on the MCDI and Vineland-II (N = 55) expressive language domains were significantly correlated, but performances on the MCDI and Vineland ABS (N = 74) expressive language domains were not significantly related. This finding may reflect true differences in the samples of toddlers from the separate projects, or it may have been influenced by psychometric differences between the Vineland versions (see methods section for a description of differences between versions). It is important to take this limitation into consideration when interpreting findings from performances on the Vineland.

Third, other psychometric limitations of the assessments used in this study should be acknowledged; specifically, the psychometric challenges of using these assessments with toddlers with developmental disabilities. Gleason and Coster (2012) found that while the Vineland-II demonstrates overall good correspondence with activity and participation codes from the International Classification of Functioning, Disability and Health for Children and Youth (IFC-CY) guidelines, many concepts measured by Vineland-II items overlap across domains, and many items measure multiple concepts. Therefore, they concluded that caution should be used when interpreting Vineland-II profile scores in children with specific communication, sensory and/or motor deficits that may impact scores across multiple domains directly and indirectly. The Mullen has also been critiqued for its use with very young toddlers and toddlers with disabilities. According to Bradley-Johnson (2001), some of these criticisms include low test-retest reliability coefficients (< .80) for children 25-56 months of age, and steep scoring gradients on items for children birth through 20 months, i.e. a one-point change in raw scores equates to a four to ten point change in T-scores. This point is problematic when attempting to evaluate and discriminate abilities among very young children, or children with severe disabilities whose performance falls at or below the 20 month age range, which was characteristic of many of the children who participated in this study.

4.4 Clinical implications and future research directions

Findings from this study make several meaningful contributions to the field. First, overall, many of the toddlers who participated in this study demonstrated consistent levels of performance across developmental domains, and group mean performances within and between tests were highly related. However, a substantial number of toddlers exhibited distinct strengths and weaknesses in their performance across domains. This finding supports previous literature that has found profiles of strengths and weaknesses in children with developmental disabilities, such as autism, and extends those findings to a group of children with severe spoken language delays. These profiles of abilities may influence children's individual responses to intervention. Practitioners should also be sensitive to the unique and individual needs of young children who may appear similar on the basis of their spoken language output, but may have very different needs according to their cognitive, language and adaptive functioning abilities.

Second, this study makes a meaningful contribution to considerations about the number and types of assessments needed to accurately evaluate toddlers with developmental disabilities. Our findings suggest that for children with the lowest expressive language ability, parent-reported and clinician-assessed outcomes across cognitive, language and adaptive functioning domains are not as strongly related as for children with more expressive language. This suggests that single-reporter assessment should be interpreted cautiously for children with very severe expressive language impairments, and clinicians should strive to utilize multi-method and multi-informant assessments to obtain the most accurate picture of these children. However, for children with global developmental delays who have greater expressive language skills, measures across reporters are more consistent, and therefore clinicians can feel more confident in drawing conclusions from single-informant data for these children.

Future studies should directly examine diagnostic validity in single versus multi-rater assessments in toddlers with severe language challenges, and investigate how it might relate to intervention decisions and efficacy. Future studies should also investigate how individual differences across developmental abilities relate to response to different forms of cognitive, language and adaptive functioning interventions, and whether or not early performance patterns in toddlers with developmental disabilities remain consistent or change over time.

In conclusion, this study revealed meaningful differences in a group of children with limited spoken language and developmental disabilities in 1) individual patterns of strengths and weaknesses, and 2) between types of assessment measures, underscoring previously documented psychometric and conceptual weakness of each measurement type, especially when utilized in isolation.

REFERENCES

Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology*, *12*(4–5), 269–277.

Baker, B. L., Blacher, J., Crnic, K. A., & Edelbrock, C. (2002). Behavior Problems and Parenting Stress in Families of Three-Year-Old Children With and Without Developmental Delays. *American Journal on Mental Retardation*, *107*(6), 433. http://doi.org/10.1352/0895-8017(2002)107<0433:BPAPSI>2.0.CO;2

Bayley, N., & Reuner, G. (2006). *Bayley scales of infant and toddler development: Bayley-III* (Vol. 7). Harcourt Assessment, Psych. Corporation San Antonio, Tex, USA.

Ben-Sasson, A., & Gill, S. V. (2014). Motor and language abilities from early to late toddlerhood: using formalized assessments to capture continuity and discontinuity in development. *Research in Developmental Disabilities*, *35*(7), 1425–1432.

Björn, P. M., Kakkuri, I., & Leppänen, P. H. T. (2014). The covariation between parental and expert evaluations of early language skills. *Early Child Development and Care*, *184*(5), 706–718.

Bradley-Johnson, S. (2001). Cognitive assessment for the youngest children: A critical review of tests. *Journal of Psychoeducational Assessment*, *19*(1), 19–44.

Bradley-Johnson, S., & Johnson, C. M. (2007). Infant and toddler cognitive assessment. *Psychoeducational Assessment of Preschool Children*, *4*, 325–358.

Brady, N. C., Marquis, J., Fleming, K., & McLean, L. (2004). Prelinguistic predictors of language growth in children with developmental disabilities. *Journal of Speech, Language, and Hearing Research*, *47*(3), 663–677.

Bruckner, C., Yoder, P., Stone, W., & Saylor, M. (2007). Construct validity of the MCDI-I receptive vocabulary scale can be improved: Differential item functioning between toddlers with autism spectrum disorders and typically developing infants. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1631–1638.

Burack, J. A., & Volkmar, F. R. (1992). Development of low- and high-functioning autistic children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *33*(3), 607–16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1374418

Burchinal, M. R., Roberts, J. E., Hooper, S., & Zeisel, S. A. (2000). Cumulative risk and early cognitive development: a comparison of statistical risk models. *Developmental Psychology*, *36*(6), 793.

Burns, T. G., King, T. Z., & Spencer, K. S. (2013). Mullen scales of early learning: the utility in assessing children diagnosed with autism spectrum disorders, cerebral palsy, and epilepsy. *Applied Neuropsychology: Child*, *2*(1), 33–42.

Buschmann, A., Multhauf, B., Hasselhorn, M., & Pietz, J. (2015). Long-Term Effects of a Parent-Based Language Intervention on Language Outcomes and Working Memory for Late-Talking Toddlers. *Journal of Early Intervention*, *37*(3), 175–189. http://doi.org/10.1177/1053815115609384

Calandrella, A. M., & Wilcox, M. J. (2000). Predicting language outcomes for young prelinguistic children with developmental delay. *Journal of Speech, Language, and Hearing Research*, *43*(5), 1061–1071.

Carpentieri, S., & Morgan, S. B. (1996). Adaptive and intellectual functioning in autistic and nonautistic retarded children. *Journal of Autism and Developmental Disorders*, *26*(6), 611–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8986847

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., … Schopler, E.

(1998). The Vineland Adaptive Behavior Scales: supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders*, *28*(4), 287–302. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9711485

Caselli, M. C., Vicari, S., Longobardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1125–1135.

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech–sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proceedings of the National Academy of Sciences*, *100*(9), 5567–5572.

Chapman, R. S. (1997). Language development in children and adolescents with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *3*(4), 307–312. http://doi.org/10.1002/(SICI)1098-2779(1997)3:4<307::AID-MRDD5>3.0.CO;2-K

Chapman, R. S., & Hesketh, L. J. (2000). l. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*(2), 84–95.

Charman, T. (2004). Matching preschool children with autism spectrum disorders and comparison children for language ability: Methodological challenges. *Journal of Autism and Developmental Disorders*, *34*(1), 59–64.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, *30*(1), 213–236.

Ciccone, N., Hennessey, N., & Stokes, S. F. (n.d.). Community-based early intervention for language delay: a preliminary investigation. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *47*(4), 467–70. http://doi.org/10.1111/j.1460-6984.2012.00149.x

Cirrin, F. M., & Rowland, C. M. (1985). Communicative assessment of nonverbal youths with severe/profound mental retardation. *Mental Retardation*, *23*(2), 52–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3990553

Coaley, K. (2014). *An introduction to psychological assessment and psychometrics*. Sage.

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech, Language, and Hearing Research*, *34*(3), 565–571.

DeVeney, S. L., Hoffman, L., & Cress, C. J. (2012). Communication-based assessment of developmental age for young children with developmental disabilities. *Journal of Speech, Language, and Hearing Research : JSLHR*, *55*(3), 695–709. http://doi.org/10.1044/1092-4388(2011/10-0148)

Disabilities, A. A. on I. and D. (2010). Intellectual disability: Definition, classification, and systems of supports. *AAIDD*.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service Circle Pines, MN.

Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: Words and Gestures, Desktop Scannable*. Brookes.

Fenton, G., D'ardia, C., Valente, D., Del Vecchio, I., Fabrizi, A., & Bernabei, P. (2003). Vineland adaptive behavior profiles in children with autism and moderate to severe developmental delay. *Autism*, *7*(3), 269–287.

Fidler, D. J., Hepburn, S., & Rogers, S. (2006). Early learning and adaptive behaviour in toddlers with Down syndrome: evidence for an emerging behavioural phenotype? *Down's Syndrome, Research and Practice : The Journal of the Sarah Duffen Centre*, *9*(3), 37–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16869373

Fidler, D. J., Philofsky, A., & Hepburn, S. L. (2007). Language phenotypes and intervention planning: Bridging research and practice. *Mental Retardation & Developmental Disabilities Research Reviews*, *13*(1), 47. http://doi.org/10.1002/mrdd.20132

Floyd, R. G., Shands, E. I., Alfonso, V. C., Phillips, J. F., Autry, B. K., Mosteller, J. A., … Irby, S. (2015). A Systematic Review and Psychometric Evaluation of Adaptive Behavior Scales and Recommendations for Practice. *Journal of Applied School Psychology*, *31*(1), 83–113. http://doi.org/10.1080/15377903.2014.979384

Fodstad, J. C., Matson, J. L., Hess, J., & Neal, D. (2009). Social and communication behaviours in infants and toddlers with autism and pervasive developmental disorder-not otherwise specified. *Dev Neurorehabil*, *12*(3), 152–157. http://doi.org/10.1080/17518420902936748

Fowler, A. E. (1990). Language abilities in children with Down syndrome: evidence for a specific syntactic delay. In D. Cicchetti & M. Beeghly (Eds.), *Children with Down syndrome: A developmental perspective* (pp. 302–328). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511581786.010

Frost, K. M., Hong, N., & Lord, C. (2017). Correlates of Adaptive Functioning in Minimally Verbal Children With Autism Spectrum Disorder. *American Journal on Intellectual and Developmental Disabilities*, *122*(1), 1–10. http://doi.org/10.1352/1944-7558-122.1.1

Gerken, K. C., Eliason, M. J., & Arthur, C. R. (1994). The assessment of at-risk infants and toddlers with the Bayley Mental Scale and the Battelle Developmental Inventory: Beyond the data. *Psychology in the Schools*, *31*(3), 181–187. http://doi.org/10.1002/1520-

6807(199407)31:3<181::AID-PITS2310310302>3.0.CO;2-Q

Gleason, K., & Coster, W. (2012). An ICF-CY-based content analysis of the Vineland Adaptive Behavior Scales-II. *Journal of Intellectual and Developmental Disability*, *37*(4), 285–293.

Harrison, P., & Oakland, T. (2003). Adaptive behavior assessment system (ABAS-II). *San Antonio, TX: The Psychological Corporation*.

Hauser-Cram, P., & Woodman, A. C. (2016). Trajectories of internalizing and externalizing behavior problems in children with developmental disabilities. *Journal of Abnormal Child Psychology*, *44*(4), 811–821.

Hedrick, D. L., Prather, E. M., & Tobin, A. R. (1984). *Sequenced inventory of communication development*. University of Washington Press Seattle, WA.

Hedrick, D. L., Prather, E. M., & Tobin, A. R. (2002). *Sequence Inventory of Comminication Development* (Revisd Edi). Los Angeles, CA: Western Psychological Services.

Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, *36*(2), 149–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11344592

Hill, E. L. (2010). Motor difficulties in specific language impairment: evidence for the Iverson account? - a commentary on Iverson's "Developing language in a developing body: the relationship between motor development and language development". *Journal of Child Language*, *37*(2), 287–92. http://doi.org/10.1017/S0305000909990444

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, *37*(2), 229–261.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Wiley Online Library.

Keller, K. M., & Fox, R. A. (2009). Toddlers with developmental delays and challenging behaviors. *Early Child Development and Care*, *179*(1), 87–92. http://doi.org/10.1080/03004430600987126

Krakow, J. B., & Kopp, C. B. (1983). The Effects of Developmental Delay on Sustained Attention in Young Children. *Child Development*, *54*(5), 1143. http://doi.org/10.2307/1129670

Laws, G., & Hall, A. (2014). Early hearing loss and language abilities in children with Down syndrome. *International Journal of Language & Communication Disorders*, *49*(3), 333–42. http://doi.org/10.1111/1460-6984.12077

Liss, M., Harel, B., Fein, D., Allen, D., Dunn, M., Feinstein, C., … Rapin, I. (2001). Predictors and correlates of adaptive functioning in children with developmental disorders. *Journal of Autism and Developmental Disorders*, *31*(2), 219–230.

LORD, C., & PICKLES, A. (1996). Language Level and Nonverbal Social-Communicative Behaviors in Autistic and Language-Delayed Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(11), 1542–1550. http://doi.org/10.1097/00004583-199611000-00024

Luyster, R. J., Seery, A., Talbott, M. R., & Helen, T.-F. (2011). IDENTIFYING EARLY-RISK MARKERS AND DEVELOPMENTAL TRAJECTORIES FOR LANGUAGE IMPAIRMENT IN NEURODEVELOPMENTAL DISORDERS. *Developmental Disabilities Research Reviews*, *17*(2), 151–159. http://doi.org/10.1002/ddrr.1109

Lyytinen, P., Laakso, M., Poikkeus, A., & Rita, N. (1999). The development and predictive relations of play and language across the second year. *Scandinavian Journal of Psychology*, *40*(3), 177–186.

Magiati, I., & Howlin, P. (2001). Monitoring the progress of preschool children with autism enrolled in early intervention programmes: problems in cognitive assessment. *Autism : The International Journal of Research and Practice*, *5*(4), 399–406. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11777256

Mervis, C. B., & Robinson, B. F. (1999). Methodological Issues in Cross-Syndrome Comparisons: Matching Procedures, Sensitivity (Se), and Specificity (Sp). *Monographs of the Society for Research in Child Development*, *64*(1), 115–130. http://doi.org/10.1111/1540-5834.00011

Mervis, C. B., Robinson, B. F., Levy, Y., & Schaeffer, J. (2003). Methodological issues in cross-group comparisons of language and cognitive development. *Language Competence across Populations: Toward a Definition of Specific Language Impairment*, 233–258.

Miller, J. F., & Miller, J. F. (1999). Profiles of language development in children with Down syndrome. *Improving the Communication of People with Down Syndrome*, 11–39.

Miller, J. F., Sedey, A. L., & Miolo, G. (1995). Validity of parent report measures of vocabulary development for children with Down syndrome. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1037–1044.

Milne, S., & McDonald, J. (2015). Assessing adaptive functioning in preschoolers referred for diagnosis of developmental disabilities. *Infants & Young Children*, *28*(3), 248–261.

Mullen, E. M. (1995). *Mullen scales of early learning*. AGS Circle Pines, MN.

Newborg, J., Stock, J., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). Batelle Developmental Inventory: Examiner's manual. *Dallas: DLM/Teacher Resources*.

Nordahl-Hansen, A., Kaale, A., & Ulvund, S. E. (2014). Language assessment in children with autism spectrum disorder: Concurrent validity between report-based assessments and direct tests. *Research in Autism Spectrum Disorders*, *8*(9), 1100–1106.

O'Toole, C., & Fletcher, P. (2010). Validity of a parent report instrument for Irish-speaking toddlers. *First Language*, *30*(2), 199–217.

Palisano, R. J., Walter, S. D., Russell, D. J., Rosenbaum, P. L., Gémus, M., Galuppi, B. E., & Cunningham, L. (2001). Gross motor function of children with down syndrome: Creation of motor growth curves. *Archives of Physical Medicine and Rehabilitation*, *82*(4), 494–500. http://doi.org/10.1053/APMR.2001.21956

Pereira, K., Basso, R. P., Lindquist, A. R. R., Silva, L. G. P. da, & Tudella, E. (2013). Infants with Down syndrome: Percentage and age for acquisition of gross motor skills. *Research in Developmental Disabilities*, *34*(3), 894–901. http://doi.org/10.1016/J.RIDD.2012.11.021

Plesa Skwerer, D., Jordan, S. E., Brukilacchio, B. H., & Tager-Flusberg, H. (2016). Comparing methods for assessing receptive language skills in minimally verbal children and adolescents with autism spectrum disorders. *Autism*, *20*(5), 591–604. http://doi.org/10.1177/1362361315600146

Provost, B., Lopez, B. R., & Heimerl, S. (2007). A comparison of motor delays in young children: autism spectrum disorder, developmental delay, and developmental concerns. *Journal of Autism and Developmental Disorders*, *37*(2), 321–328.

Ray-Subramanian, C. E., Huai, N., & Ellis Weismer, S. (2011). Brief report: adaptive behavior and cognitive skills for toddlers on the autism spectrum. *Journal of Autism and Developmental Disorders*, *41*(5), 679–84. http://doi.org/10.1007/s10803-010-1083-y

Rechetnikov, R. P., & Maitra, K. (2009). Motor impairments in children associated with impairments of speech or language: a meta-analytic review of research literature. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association*, *63*(3), 255–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19522134

Ross, G., & Weinberg, S. (2006). Is there a relationship between language delays and behavior and socialization problems in toddlers? *Journal of Early Childhood and Infant Psychology*, *2*, 101–116. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2007-14172-008&site=ehost-live&scope=site

Sattler, J. M., & Hoge, R. D. (2006). Assessment of children: Behavioral, social, and clinical foundations. *USA: Publisher Inc*.

Scattone, D., Raggio, D. J., & May, W. (2011). Comparison of the Vineland Adaptive Behavior Scales, Second Edition, and the Bayley Scales of Infant and Toddler Development, Third Edition. *Psychological Reports*, *109*(2), 626–34. http://doi.org/10.2466/03.10.PR0.109.5.626-634

Seol, K. I., Song, S. H., Kim, K. L., Oh, S. T., Kim, Y. T., Im, W. Y., … Cheon, K.-A. (2014). A comparison of receptive-expressive language profiles between toddlers with autism spectrum disorder and developmental language delay. *Yonsei Medical Journal*, *55*(6), 1721–1728.

Shevell, M., Majnemer, A., Platt, R. W., Webster, R., & Birnbaum, R. (2005). Developmental and Functional Outcomes at School Age of Preschool Children With Global Developmental Delay. *Journal of Child Neurology*, *20*(8), 648–654. http://doi.org/10.1177/08830738050200080301

Singer Harris, N. G., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with williams and down syndromes. *Developmental Neuropsychology*, *13*(3), 345–370. http://doi.org/10.1080/87565649709540683

Sparrow, S. S., Balla, D. A., Cicchetti, D. V, & Doll, E. A. (1984). *Vineland adaptive behavior scales: Survey form manual*. American Guidance Service.

Sparrow, S. S., Cicchetti, D. V, & Balla, D. A. (2005). *Vineland adaptive behavior scales* (2nd editio). Circle Pines, MN: American Guidance Service.

Sullivan, J. R., Winter, S. M., Sass, D. A., & Svenkerud, N. (2014). Assessing Growth in Young Children: A Comparison of Raw, Age-Equivalent, and Standard Scores Using the Peabody Picture Vocabulary Test. *Journal of Research in Childhood Education*, *28*(2), 277–291. http://doi.org/10.1080/02568543.2014.883453

Swineford, L. B., Guthrie, W., & Thurm, A. (2015). Convergent and divergent validity of the Mullen Scales of Early Learning in young children with and without autism spectrum disorder.

Tager-Flusberg, H. (2000). The challenge of studying language development in children with autism. In L. Menn, N. Bernstein Ratner, L. (Ed) Menn, & N. (Ed) Bernstein Ratner (Eds.), *Methods for studying language production.* (pp. 313–332). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1999-16036-014&site=ehost-live&scope=site

Thal, D. J., O'Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a Parent Report Measure of Vocabulary and Syntax for Preschool Children With Language Impairment. *Journal of Speech Language and Hearing Research*, *42*(2), 482.

http://doi.org/10.1044/jslhr.4202.482

Waterhouse, L., & Fein, D. (1982). Language skills in developmentally disabled children. *Brain and Language*, *15*(2), 307–333. http://doi.org/10.1016/0093-934X(82)90062-1

Weismer, S. E., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders*, *40*(10), 1259–1273.

Wilkins, J., & Matson, J. L. (2009). A comparison of social skills profiles in intellectually disabled adults with and without ASD. *Behav Modif*, *33*(2), 143–155. http://doi.org/10.1177/0145445508321880