3-31-2020

MacArthur-Bates Communicative Developmental Inventories (CDI): A Research Synthesis Evaluating Children at 2-36 months

Nicholas Giammarco
Florida International University, ngiam002@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

MACARTHUR-BATES COMMUNICATIVE DEVELOPMENTAL INVENTORIES (CDI): A RESEARCH SYNTHESIS EVALUATING CHILDREN AT 2-36 MONTHS

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF ARTS

in

LINGUISTICS

by

Nicholas Giammarco

2020

To: Director, Linguistics Program
College of Arts, Sciences and Education

This MA Project, written by Nicholas Giammarco, and entitled, MacArthur-Bates Communicative Developmental Inventories (CDI): A research synthesis evaluating children at 2-36 months, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this MA Project and recommend that it be approved.

_____
Tometro Hopkins

_____
Ellen Thompson

_____
Melissa Baralt, Major Professor

Date of Defense: 3/31/20

The MA Project of Nicholas Giammarco is approved.

_____
Prof. Ellen Thompson
Director, Linguistics
College of Arts, Sciences, and Education

Florida International University, 2020

ABSTRACT OF THE THESIS

MACARTHUR-BATES COMMUNICATIVE DEVELOPMENTAL INVENTORIES (CDI): A RESEARCH SYNTHESIS EVALUATING CHILDREN AT 2-36 MONTHS

by

Nicholas Giammarco

Florida International University, 2020

Miami, Florida

Professor Melissa Baralt, Major Professor

This synthesis will touch on the current parent-based assessments available while focusing specifically on the MacArthur-Bates Communicative Inventory English and Spanish versions. It will analyze studies that have used this test to predict language delays in infants from 2-36 months and look at its validity and effectiveness. It will use the PRISMA method to narrow search results. The PRISMA method is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses. Areas of concern were socioeconomic status, level of parent education, race, design, and effects of disability on CDI performance. 26 studies met the criteria to be used in this synthesis. The main aspect targeted was data on vocabulary production. Results showed just under half of the CDI scores were compromised by one or more of the variables analyzed. This study concluded results of the CDI should be reviewed with caution.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

LIST OF ABBREVIATIONS AND ACRONYMS

ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
CDI	Communicative Developmental Inventories
DAY-2	Developmental Assessement of young children
ECI	Early Communicative Scale
ERIC	Online Library of Education and Inforamtion
F	Cumulative distribution function
n	Number of studies
p	Probability
PDD-NOS	Pervasive Developmental Disorder- Not otherwise specified
PLS	Preschool Language Scale
PPVT	Peabody One Word Picture Test
PRISMA	Preferred Reporting Items for Systematic Review
QUILS	Ouick Interactive Language Screener
r	Coefficient of correlation
SES	Socioeconomic status
SLPS-4	Speech-language pathologists
t	t-statistic

I. INTRODUCTION

Language development occurs for humans at birth and is a crucial process in the way we perceive our world (Colona, 2019). It is arguably the quintessential factor that separates humans from other forms of species. Many different aspects of human life can affect language development in children and infants. Interactions with family members, teachers and even other peers can influence how a child's language skills progress or regress (Colona, 2019). Literature can influence a child's language development as well. Children and infants can experience this through reading to themselves or having someone read to them like a parent, teacher or sibling (Colona, 2019).

Psychological and physiological factors may also influence the ability of a child to develop language (Colona, 2019). Disorders involving a child's hearing or speech such as a disfluency condition like stuttering all can have adverse effects on language development. These issues can result in problems at school for the child as well.

When we break down language into individual parts, there are three main areas. Phonology, Semantics, and Syntax. Phonology describes the rules that are in place to interpret sounds. Semantics involves rules that help people interpret the meaning of words. Syntax refers to the combining of words to form sentences (Colona, 2019). It is the combination of these elements that children use to develop language.

There are various ways to test a child to make sure they are on the right track with their language development skills. The most accurate way is to take them to a clinician so that the child can be professionally assessed. Speech-language pathologists are a common resource that parents have at their disposal when looking for a clinical assessment of their child's language skills. However, speech-language pathologists can be expensive and require the child to spend an

extended amount of time in a clinical setting that may cause them to lose focus. Some children do not interact well with people they do not know making assessment by a clinician extremely difficult (Feldman, et al., 2000). This is where a parent-based language assessment can come in handy. These assessments are not interventions themselves, but they can be used as tools to help decide about future intervention.

Although there are studies on the effectiveness of the MacArthur-Bates Communicative Inventories, few syntheses have been done on the Spanish and English versions with toddlers 2 - 36 months. This information is necessary to determine if the results of the MacArthur-Bates CDI are useful in predicting future language delay in children in this age group. A study by Feldman looked at the effectiveness of the MacArthur-Bates CDI with monolingual English toddlers from 1-2 years of age (Feldman, et al., 2000).

This synthesis will attempt to expand on the current studies in a few ways. First, a list of all the current parent based and language assessment tools are listed that can be used on Spanish and English-speaking toddlers from 2-36 months. These assessments are described, their reliability is shown, and their cost is given. Second, the MacArthur-Bates CDI is reviewed on its effectiveness with both English and Spanish speaking children from 2-36 months of age. Effects of socioeconomic status, education, race, design, and disability will be analyzed and reported on. There were three main objectives this study focused on. The first was to analyze any effects that socioeconomic status, education, race, design, and disability had on the effectiveness of testing outcomes. Second, to show the effectiveness of the MacArthur-Bates CDI on both English and Spanish speaking children from 2-36 months of age. Finally, ways to edit the MacArthur-Bates CDI were discussed.

II. REVIEW OF LITERATURE

2.1 Home-based parental assessments

There is another way that language acquisition can be tested in children. Parent based language assessments have been available to families to make the task more convenient and less financially burdensome. These parent based assessments allow for children to be analyzed at home where they can have multiple observations over a longer interval of time (Feldman, et al., 2000). Also, parent-based assessments can be completed without the child's complete cooperation which is a helpful advantage for parents (Feldman, et al., 2000).

There are many types of parent-based assessments available today. The following are a list of current parent-based assessments: Peabody One Word Picture Test, Bayley's Language Scale, Expressive One Word Vocabulary Test, The Battelle Developmental Inventory (BDI), Vineland Adaptive Behavior Scale, The Brigance Diagnostic Inventory of Early Development. Communication and Symbolic Behavior Scales, The Developmental Assessment of Young Children, Merrill-Palmer revised scale of development, Mullen Scales of Early Learning, PCI, Early Communication Indicator, Quick Interactive Language Screener, and the MacArthur-Bates CDI.

2.2 Peabody One Word Picture Test

The Peabody One Word Picture Test (PPVT) is an assessment of vocabulary that can be used on children as young as 30 months through adulthood (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018). The assessment has been shown to have good reliability and validity statistics. The participants are given a word and are then asked to find the corresponding picture that matches (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018). A study looked at 49 monolingual English children were given the PPVT at 48 months of age (Friend, Smolak, Liu,

Poulin-Dubois, & Zesiger, 2018). The vocabulary scores ranged from 31 to 106 words, the mean was 72.73 and the Standard Deviation was 15.48 (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018).

The range was between the 13th to the 99th percentile (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018). This was good enough for an Internal Consistency of .94 percent (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018). The test was used on children from 18-24 months in another study (Fletcher, 2005). The study looked at 25 children from low income home environments to assess responsiveness while completing picture book reading (Fletcher, 2005). There was an increase in responsiveness from the 3rd and 4th assessment while joint attention increased only during assessments 1 and 2 (Fletcher, 2005). More research was deemed necessary when determining the validity of the Peabody for children under 30 months (Fletcher, 2005).

2.2.1 Bayley's Language Scale

The Bayley's Language Scale is an assessment that is used to interpret the developmental abilities of toddlers from 1 to 42 months of age (Armstrong & Agazzi, 2010). It looks at the following parts of development, cognitive, language, motor, adaptive, and social-emotional development (Armstrong & Agazzi, 2010). The assessment is very helpful at determining toddler and child processing speed, which Armstrong defines as the ability to complete new tasks correctly (Armstrong & Agazzi, 2010). This is monitored by having a child complete tasks that include completion of puzzles that increase in difficulty as the assessment progresses (Armstrong & Agazzi, 2010).

The advantage of the Bayley's Language Scale is its focus on making the application of the measures as simple as possible to promote more interaction with the child (Armstrong & Agazzi,

2010). Some have posited that the Bayley's is not the best assessment for cognitive skills because some of the elements on the test have little educational bearing (Armstrong & Agazzi, 2010). However, it has widely been categorized as the premier test for assessing infant and toddler development (Armstrong & Agazzi, 2010).

2.2.2 Expressive One Word Vocabulary Test

The Expressive One Word Vocabulary Test was created in 1979 to help with addressing potential developmental issues with speech defects, learning disorders, bilingual fluency in English, auditory processing, and auditory-visual processing ability (Canestraro, 2014). It was intended to be used with children between the ages of 12-16 (Canestraro, 2014). The test consists of 70 pictures that are displayed in relation to their difficulty (Canestraro, 2014). The pictures vary from one tangible object to multiple objects that incorporate conceptual ideas (Canestraro, 2014). The pictures that are used in the test are commonly known and the tests avoids images that relate to a certain sex, race, or culture (Canestraro, 2014).

A solid mark for test reliability around is 0.90 or higher when looking at an instrument's effectiveness (Canestraro, 2014). The reliability of the Expressive One Word Vocabulary Test is between 0.89 and 0.94 percent with an average of 0.92 which gives this assessment a very reliable rating (Canestraro, 2014). The test was used on 24 toddlers at 2 years of age to test its validity in a study by Dale (Dale, 1991). The results showed high validity on the assessment with this age range with vocabulary scores and syntactic development outcomes (Dale, 1991).

2.2.3 The Battelle Developmental Inventory

The Battelle Development Inventory interprets five aspects of child development which include personal-social, adaptive, motor, communication, and cognition (Johnson, Cook, & Kullman, 1992). The test is used to access English and Spanish speaking children from birth to 8

years old and generally takes about 1-2 hours to complete (Johnson, Cook, & Kullman, 1992). The Battelle Development Inventory provides a total score at the end of the assessment as well as total scores for each of the five areas of development it measures (Johnson, Cook, & Kullman, 1992). The test uses t-scores, percentile ranks, age equivalents, z-scores, and normal curve equivalents (Johnson, Cook, & Kullman, 1992).

The test uses information taken from several sources including, interviews with caregivers, observations, and structured assessments (Johnson, Cook, & Kullman, 1992). The test can be modified for children who have certain disabilities including hearing, visual, and motor (Johnson, Cook, & Kullman, 1992). The Battelle Development Inventory has generally been known to have high reliability compared to other assessments, however it can be costly at around $900 (Johnson, Cook, & Kullman, 1992).

2.2.4 Vineland Adaptive Behavior Scale

The Vineland Adaptive Behavior Scale is used for English and Spanish speaking children from birth to 18 years of age and takes around 20-60 minutes to complete (Johnson, Cook, & Kullman, 1992). The information is obtained through interviews with caregivers and reveals data on four aspects of development such as, communication, daily living skills, socialization, and motor skills (Johnson, Cook, & Kullman, 1992). The Vineland Adaptive Behavior Scale yields an Adaptive Behavior Composite score which involves transforming raw scores into standard scores, national percentile ranks, stanines, adaptive levels, and age equivalents (Johnson, Cook, & Kullman, 1992). The Vineland Adaptive Behavior Scale has been shown to have high reliability and is relatively inexpensive at around $230 (Johnson, Cook, & Kullman, 1992).

2.2.5 The Brigance Diagnostic Inventory of Early Development

The Brigance Diagnostic Inventory of Early Development assesses 98 abilities in 11 areas of development that include, perambulatory, gross, and fine motor skills; self-help skills; speech and language skills; general knowledge and comprehension; social and emotional development; readiness; basic reading skills; manuscript writing; and basic math (Holahan & Costenbader, 2000). The test can be used on English and Spanish speaking children from birth to 7 years of age and takes about 25-30 minutes to administer (Holahan & Costenbader, 2000).

The test uses developmental age as a determiner for how the skills tested on are structured and can be used to create educational goals (Holahan & Costenbader, 2000). Data on the reliability and validity of the Brigance Diagnostic Inventory of Early Development is not as available as other assessments, but the test is known for its adaptability and organizational ability (Holahan & Costenbader, 2000). It is also extremely affordable compared to other assessments at $185 (Holahan & Costenbader, 2000).

2.2.6 Communication and Symbolic Behavior Scales

The Communication and Symbolic Behavior Scales was engineered to record early communication and behavior skills in English speaking infants and young children from 6 months to 2 years of age (Eadie, et al., 2010). The test was designed to help with locating communication delays by using a Behavior Sample and a Caregiver Questionnaire (Eadie, et al., 2010). The Communication and Symbolic Behavior Scales measure three areas of communication including, social, speech, and symbolic (Eadie, et al., 2010). The social component looks at eye gaze, gesture, and joint attention (Eadie, et al., 2010). The speech component incorporates consonant inventory and consonant vowel combinations in syllables or words (Eadie, et al., 2010). The symbolic component looks at early recognition of response to

name and body parts (Eadie, et al., 2010). The test has been shown to have high reliability, but it is expensive at around $399 (Eadie, et al., 2010).

2.2.7 Developmental Assessment of Young Children

The Developmental Assessment of Young Children (DAYC-2) was created for use on English speaking children from birth to five years of age (Judith & Maddux, 2014). The DAYC-2 was designed to look at five areas of development including, cognitive, communication, social-emotional, physical development, and adaptive behavior (Judith & Maddux, 2014). The goal is to detect potential language delays in children in any of these five areas so that intervention can be implemented to promote improvement (Judith & Maddux, 2014). A child may be tested on all areas and obtain what is called the General Development Index (GDI), or only the domains that are of relevance to the child's current needs (Judith & Maddux, 2014).

Each part of the DAYC-2 takes between 10-20 minutes to complete which is quick compared to other assessments, however the assessment can run between 50-100 minutes if all areas are tested (Judith & Maddux, 2014). The results are recorded as quotients (M=100, SD=15), age equivalents, or percentiles (Judith & Maddux, 2014). Results can be acquired through observing a child, talking with a caregiver, direct assessment, or any mixture of these means (Judith & Maddux, 2014). This assessment has been shown to be reliable with scores falling from 0.89 to 0.98 for the domains and 0.82 to 0.97 for the subdomains (Judith & Maddux, 2014). 0.90 is considered the proper reliability score for an effective assessment (Judith & Maddux, 2014). The test can cost around $235 (Judith & Maddux, 2014).

2.2.8 Merrill Palmer Revised Scale of Development

The Merrill Palmer Revised Scale of Development is an assessment used for English and Spanish speaking children ages 1 month to 6 years 6 months of age to identify developmental

delays and recommendations for early treatment (Roid & Sampers, 2004). The instrument records data on five language areas including, cognitive, language, motor, self-help, and social-emotional (Roid & Sampers, 2004). The assessment consists of examiner forms and parent reports and contains four assessment batteries (Roid & Sampers, 2004). There is a cognitive battery which looks at, general cognitive, receptive language, and fine motor areas, supplemental scores for memory, speed of cognition, and visual motor ability (Roid & Sampers, 2004).

It contains a gross motor scale which includes information on general gross motor development, unusual movements, and atypical movement patterns (Roid & Sampers, 2004). The test has four Social Emotional Scales which are the Examiner Observation Form/Test-Session Behavior, the Social-Emotional Developmental Scale-Parent Report, the Social-Emotional Temperament Scale-Parent Report, and the Social-Emotional Problem Indicators (Roid & Sampers, 2004). The raw scores can be converted into standard scores, percentile ranks, age equivalents, and growth scores (Roid & Sampers, 2004). The assessment can take around 45 min to administer and has a high reliability exceeding 0.90, however, it is costly at around $1,070 (Roid & Sampers, 2004).

2.2.9 Mullen Scales of Early Learning

The Mullen Scales of Early Learning can be used to assess children from birth to 5 years and 8 months of age (Mullen, 1995). The test provides data on five domains of language including Gross Motor, Expressive Language, Visual Reception, Receptive Language, and Fine Motor (Mullen, 1995). The time it takes to assess a child can vary on their age, but anywhere from 15-60 minutes is an accurate timeframe for completion (Mullen, 1995). Each of the test areas are recorded with T-scores, percentile, and age equivalent scores (Mullen, 1995).

A physical printed score can be obtained where the child's performance can be easily observed, however, a computer score can be chosen instead (Mullen, 1995). The computer report will convert all the raw scores and give a treatment plan which includes activities the parent can engage with their child in at home (Mullen, 1995). The assessment has a high reliability rating and costs around $728 (Mullen, 1995).

2.2.10 Preschool Language Scale PLS

The Preschool Language Scale is a assessment that is designed to interpret a English and Spanish speaking child's auditory and expressive language skills from ages birth to 6 years 11 months (H. Qi & Marley, 2011). The PLS is broken up into two scales, the Auditory Comprehension Scale (AC) and the Expressive Language Scale (EL) to uncover possible language delays and to monitor language delays over time (H. Qi & Marley, 2011). The PLS can take anywhere from 20-45 min to administer and has been shown to have a high reliability and internal consistency rating (H. Qi & Marley, 2011).

2.2.11 Early Communication Indicator

The Early Communication Indicator (ECI) is an interactive language assessment test used to track changes in a toddler's expressive communication skills (Greenwood, 2010). It is a 6-minute activity focused test can be given every few months or even once a month to check on a child's language development process (Greenwood, 2010). It can be used on children starting at 6 months and uses progress monitoring measures to document a child's outcomes during early educational experiences (Greenwood, 2010). The ECI is given by an adult that the child has a familiar relationship with (Greenwood, 2010). The adult engages the child in playful activity with the Fischer-Price Barn (Form A) or Fischer-Price House (Form B) (Greenwood, 2010).

The test usually will commence in a setting that is familiar to the child, either the home or a school like setting where distractions are limited (Greenwood, 2010). Sometimes the test can be given by two adults where one is interacting directly with the child and another is recording the results (Greenwood, 2010). The reliability of this assessment was validated when a total of 246 tests were done and analyzed (Greenwood, 2010). The results showed that the reliability of the outcomes was high coming in at .96 (Greenwood, 2010). A 2010 study by Greenwood showed reliability of the assessment in toddlers as early as 6 months (Greenwood, 2010).

2.2.12 Quick Interactive Language Screener (QUILS)

The Quick Interactive Language Screener (QUILS) is an assessment with roots in language acquisition (Levine, 2020). There are 12 subsets of language analyzed covering three language areas including vocabulary, syntax, and the language learning process (Levine, 2020). This assessment can by performed by parents and other nonprofessionals and takes around 15 minutes to complete (Levine, 2020). QUILS can be used on children from 3-5 years of age and is a dialect neutral test to discourage any language biases (Levine, 2020). The vocabulary portion looks at a child's knowledge of open class words (Nouns, Verbs) and closed class words (Prepositions, Conjunctions) (Levine, 2020). The syntax portion assess a child's understanding of different syntactic structures that include sentences looking at past actions and locations, sentences with multiple modifiers, embedded clauses, and wh-questions (Levine, 2020).

The language portion looks at how children analyze new words and understand syntactic structures to brand new words (Levine, 2020). The reliability of QUILS was assessed in a study that incorporated 674 preschool age children in Head Start Programs in five states (Levine, 2020). The assessment had a reliability of .93 which is very high (Levine, 2020). In a study by Levine (2020), and colleagues, the QUILS was used to assess 3, 4, and 5-year-old children on

vocabulary, syntax, and language abilities from low and middle-income families. The research found that SES played a crucial role in the child's language outcomes and the difference between low- and middle-income family results were large (Levine, 2020). The researchers mentioned that mentioned that looking at factors beyond just vocabulary and syntax to explain the SES difference in language (Levine, 2020).

2.3 MacArthur-Bates CDI

The MacArthur-Bates Communicative Inventory English version is a parent-based assessment that asks what words a child can produce and what words a child understands (Pearson, 1994). It contains lists of vocabulary words that parents can use to assess their child's knowledge of through at home testing (Pearson, 1994). The assessment is relatively inexpensive compared to other language assessments at around $100. It can take 20-40 minutes to complete. It attempts to compare the language abilities of typical and atypical children to respective language norms to predict potential issues of language delay (Pearson, 1994). The MacArthur-Bates Inventario del Desarrollo de las Habilidades Spanish version is not a direct translation of the English version (Pearson, 1994). It does have around an 80% overlap between the English version (Pearson, 1994). It also reflects linguistic and cultural aspects that are different than those found on the English version (Pearson, 1994).

Gulberson (2011), and partners looked at the short form Spanish version of the CDI and its validity in measuring developmental delays in monolingual Spanish toddlers with down syndrome. The results showed a strong validity for the CDI with Spanish speaking toddlers with down syndrome from 24 to 35 months of age (Gulberson, 2011). The only criticism was that the short forms of the Spanish CDI did not ask for the three longest utterances so that the mean length of utterance could not be computed (Gulberson, 2011).

Heilmann (2005), and colleagues looked at the validity of the MacArthur-Bates CDI with monolingual English speaking late-talking toddlers at 30 months of age. Results showed that at 24 months late talking toddlers had scores below the 10th percentile on productive vocabulary, however scores rose to above the 15th percentile at 30 months (Heilmann, 2005). This data was in favor of the utility of the CDI in predicting language out outcomes in late talking toddlers (Heilmann, 2005).

De Diego-Lazaro (2019), and colleagues studied how monolingual Spanish speaking toddlers who are deaf or hard of hearing acquire vocabulary by using the MacArthur-Bates CDI. There was a total of 53 participants between the ages of 8-34 months (de Diego-Lazaro, 2019). The study focused on how maternal education effected the vocabulary scores on the CDI (de Diego-Lazaro, 2019). They found that maternal education was not a significant factor in predicting vocabulary development (de Diego-Lazaro, 2019).

Houston-Price (2007), and colleagues looked at how parents reported their toddler's language outcomes using the MacArthur-Bates CDI. The participants were monolingual English speakers at 1 year and 3 months, 1 year and 6 months, and 1 year and 9 months (Houston-Price, 2007). There were two separate studies which looked at how parents recorded their child's knowledge of known and unknown words using the CDI (Houston-Price, 2007). After the CDI, the infants participated in a laboratory task to compare results (Houston-Price, 2007). In both studies, the researchers discovered that parents reported their children comprehended many unknown words and suspected that they may be overestimating their vocabulary abilities (Houston-Price, 2007). Both studies were rife with false negatives in infant vocabulary production as the rate of comprehension of known words and unknown words was relatively the same (Houston-Price, 2007).

Hurtado (2014), and colleagues were able to use the results from the MacArthur-Bates CDI to show vocabulary comprehension in 37 bilingual English and Spanish speaking toddlers at 30 and 36 months. Both expressive and receptive vocabulary scores showed toddlers knew more Spanish than English words at both ages (Hurtado, 2014). It was noted that there was considerable range encompassing the sample with a few children producing a lot more Spanish words than English words and vice versa (Hurtado, 2014).

One study found mixed results when determining whether socioeconomic status influenced monolingual Spanish speaking toddlers on the MacArthur-Bates CDI (Jackson-Maldonado. D., 2013). They broke up 601 participants into an 8-12-month-old group and a 13-18-month-old age group (Jackson-Maldonado. D., 2013). They found that in the younger age group, children from low SES backgrounds were assessed as having a higher number of words understood than children with higher SES (Jackson-Maldonado. D., 2013).

There was not a similar correlation in the older age group, however, vocabulary production showed higher scores in the older age group with no effect from SES (Jackson-Maldonado. D., 2013). Vocabulary comprehension showed lower scores for children in the younger age group that had educated families and higher in younger children from uneducated families (Jackson-Maldonado. D., 2013). The older age group showed the opposite with children from educated families scoring higher than uneducated families (Jackson-Maldonado. D., 2013).

Pearson (1994), and colleagues did a study on bilingual Spanish and English toddlers from 10-30 months using the MacArthur-Bates CDI. They determined that trying to measure each toddler's vocabulary score on the Spanish and English versions of the CDI separately led to underrepresentation in middle class bilingual results (Pearson, 1994). They thought a total conceptual vocabulary score should be used for bilingual toddlers (Pearson, 1994). This is

acquired by combining the Spanish and English scores and subtracting translation equivalents to get a more accurate representation of bilingual vocabulary output (Pearson, 1994). A translation equivalent would be a word that means the same in English as it does in Spanish (Pearson, 1994). In this case a child would not receive double credit for knowing both *gato* and *cat*. The results showed that bilingual children's separate Spanish and English CDI scores were below monolingual vocabulary norms (Pearson, 1994). However, when the total conceptual score was analyzed, scores fell within normal developmental ranges (Pearson, 1994).

Luyster (2007), and colleagues did a study on how the MacArthur-Bates CDI recorded results of 153 monolingual English toddlers with autism on vocabulary compared to toddlers with developmental delay and typical development. The autism group scored lower than the two other groups on verbal IQ, nonverbal IQ, words known, and phrases understood (Luyster, 2007). The researchers expressed that these results were comparable with published norms and proved the CDI could be used as an accurate parent based assessment tool (Luyster, 2007).

A study by Mancilla-Martinez (2011), and colleagues looked at how 79 Spanish and English bilingual toddlers 24-36 months from low-income families fared with parent reports on the MacArthur-Bates CDI. The study showed that low income parents could distinguish between words their child said and words they understood (Mancilla-Martinez, 2011). The study showed that parents had an easier time reporting information about their child's primary language rather than their secondary language (Mancilla-Martinez, 2011).

Marchman (2002), and colleagues looked at the utility of the MacArthur-Bates CDI with 26 Spanish and English bilingual toddlers from 24-36 months of age. They looked at factors like maternal education, what language was spoken at home. proportion of English to Spanish output and mother's acculturation level (Marchman, 2002). The reported and observed language

measures were solid even after taking these variables into consideration (Marchman, 2002). Similar strong connections were also maintained when measuring grammar and mean length of utterance (Marchman, 2002). Results also had strong correlations when looking at whether the forms were done by either one or multiple people (Marchman, 2002). Even speakers that spoke both English and Spanish could correctly record the child's English and Spanish word knowledge (Marchman, 2002). The same was not true for grammar reporting as Spanish had strong correlations in the overlap group, but weaker correlations in the non-overlap group (Marchman, 2002).

McDuffie (2005), and colleagues used the MacArthur-Bates CDI to assess 29 two and three-year-old monolingual English speakers with autism. The researchers were able to record accurate results when it came to comprehension and production skills (McDuffie, 2005). They came to the conclusion the CDI was an accurate tool to use for looking at the language skills of developing toddlers with autism (McDuffie, 2005).

Scherer (1999), used the MacArthur-Bates CDI to measure the vocabulary production of three 2-year old monolingual English speakers with developmental delays. The toddlers were tested before and after a language intervention program (Scherer, 1999). The CDI showed improvement pre-test and post-test on all three areas of language measured including the number of vocabulary words, mean of the three longest sentences, and number of suffixes used (Scherer, 1999).

Thal (1999), and colleagues used the MacArthur-Bates CDI to measure language abilities of 12 monolingual English-speaking toddlers from 24-32 months. The study determined the CDI was applicable to assessing vocabulary in young toddlers with language delay, however, comprehension scores and gestures were not as valid (Thal, 1999). The number of gestures

produced did show a relationship with the Preschool Language Scale which may show that this measure can be effective with language delayed toddlers who have difficulty responding with accuracy on the behavioral dimension of language comprehension (Thal, 1999).

A research study by Pan and Rowe (2004), wanted to show how low-income families scored their monolingual English children at 2 and 3 months of age on the MacArthur-Bates CDI. They also looked at factors like maternal age, education, and race to see if they effected the scoring results in any way (Pan & Rowe, 2004). The researchers found that maternal education did not correlate with how the children performed on the CDI at 2 months of age (Pan & Rowe, 2004). They found that maternal education had more of an effect on child performance on the CDI at 3 months of age (Pan & Rowe, 2004).

The study found that maternal language and literacy skills had a greater effect on child vocabulary growth than maternal education (Pan & Rowe, 2004). Maternal age did not show much effect on child scores on the CDI at 2 months, but did show a positive correlation with receptive language at 3 months (Pan & Rowe, 2004). Race and ethnicity did have an effect on CDI scores with white mothers scoring their 2 month old's higher than black and Hispanic mothers (Pan & Rowe, 2004).

In Brady and Goodman (2014), the researchers looked at how maternal education correlated with the scores of monolingual English-speaking children at 18, 24, 30, and 36 months on the vocabulary portion of the MacArthur-Bates CDI. The results showed that the higher the maternal education the lower percentile the child scored in on all four stages of testing on vocabulary (Brady & Goodman, 2014). The mean length of utterance increased as maternal education increased (Brady & Goodman, 2014).

A different study looked at the vocabulary scores of low- and middle-income children on the MacArthur-Bates CDI (Arriaga, Fenson, Cronan, & Pethick, 1998). The children were between 16 and 30 months of age (Arriaga, Fenson, Cronan, & Pethick, 1998). There was a small negative relationship between the vocabulary production and socioeconomic status (Arriaga, Fenson, Cronan, & Pethick, 1998). The study showed that the lower-income class of children had much lower scores on the three areas tested including vocabulary production, combining words, and sentence complexity (Arriaga, Fenson, Cronan, & Pethick, 1998).

The researchers offered the ideas that lower-income parents have less language output with their children which might cause them to underestimate their child's scores (Arriaga, Fenson, Cronan, & Pethick, 1998). They also said that middle-income parents might give their children higher scores because of social desirability (Arriaga, Fenson, Cronan, & Pethick, 1998). Either way, the study further displayed how trepidation should be used when making judgements about potential language delay in a child while using the CDI (Arriaga, Fenson, Cronan, & Pethick, 1998).

One study compared the vocabulary scores of pre-term infants to that of regularly developing infants on the MacArthur-Bates CDI (Foster-Cohen, J, Champion, & Woodward, 2007). The CDI did accurately show that children born very preterm are at a elevated risk for language delays in their language development (Foster-Cohen, J, Champion, & Woodward, 2007). The study showed no between group differences in socioeconomic status or maternal education when it came to vocabulary scores (Foster-Cohen, J, Champion, & Woodward, 2007).

Another study looked at the vocabulary scores of 23 monolingual English children from low-income and middle-income families on the MacArthur-Bates CDI (Furey, 2011). The children were accessed at 16 and 18 months by their mothers (Furey, 2011). The results showed that low-

income children had lower scores on words produced than middle-income children at both 16 and 18 months (Furey, 2011). However, income levels did not affect vocabulary comprehension scores (Furey, 2011). Maternal reporting accuracy was accurate between low- and middle-income families (Furey, 2011). When the CDI were reviewed by professional clinicians, the accuracy of the maternal reports rose from the 16 month visit to the 18 month visit (Furey, 2011). The researchers posited this happened because the mothers become more familiar with the test the second time and therefore could score their children better (Furey, 2011).

Feldman (2000), and partners looked at the vocabulary scores of 2,156 monolingual English speaking toddlers between 1 and 2 years of age. They looked at the effects of maternal education as a factor in how the scores were recorded (Feldman, et al., 2000). The results showed that mothers with lower education recorded higher scores for their children than mothers with high education (Feldman, et al., 2000). The researchers thought this result may be caused by lower educated mothers overestimating their child's scores because of not understanding what the CDI directions (Feldman, et al., 2000).

A study by Fenson (1994), and partners found similar results when comparing maternal education to CDI vocabulary scores. Another study by Roberts (1999), and partners looked at vocabulary scores on the MacArthur-Bates CDI of monolingual English speaking African American children from low-income families. The children were between the ages of 13 and 18 months (Roberts, 1999). This study showed children that came from home environments that are more supportive had higher vocabulary scores, regular nouns and verb scores, and longer utterances (Roberts, 1999).

A study by Ramirez (2019), and partners looked at how parent coaching involvement effected child language ability and child language skills later in life. The researchers used

families from diverse socioeconomic backgrounds and used the MacArthur-Bates CDI to record results (Ramírez, 2019). They looked at the changes in the language abilities of the children from the 6-month mark to the 18-month mark (Ramírez, 2019). The data was analyzed and recorded using LENA audio recorders (Ramírez, 2019). The study did not show a correlation between the SES and the intervention effects (Ramírez, 2019). It did show that parent coaching did have a positive effect on child outcomes on the CDI (Ramírez, 2019).

A study by Bates (1994), and colleagues used the MacArthur-Bates CDI to analyze stylistic and developmental variation of vocabulary in children from 8 months to 1 year and 4 months of age. Results showed a large difference in the range of number of words produced from 1 year to 1 year and 4 months (Bates, 1994). There was also a large difference in productive vocabulary (Bates, 1994). The researchers mentioned that many parents had difficulty differentiating between their child understanding a word and just being able to repeat it (Bates, 1994).

It is important to note that the relationship between socioeconomic status and issues of academic disparity have been disputed in research. In an article by Johnson and Zentella (2017), the issue of the language gap is addressed. This is the idea that the low socioeconomic status of families adversely effects educational performance (Johnson E. Z., 2017). The articles stresses that languages that are linguistically diverse to English are frowned upon especially if they are originating from a lower income group of people (Johnson E. Z., 2017). Because of this emphasis on an Americanized or English way of speaking, schools are inadequately prepared to handle children that come from different racial backgrounds (Johnson E. Z., 2017). The article points out that families from low SES situations should not be the first to be blamed for poor knowledge of language skills (Johnson E. Z., 2017). The school system plays a larger role in the

case for poor language performance in low income racially diverse children (Johnson E. Z., 2017).

Another article by Johnson (2015), explained how accurately assessing a child's language abilities can only be done by understanding the cultural situation they live in. Racially diverse families from low-income homes are often held to an unfair standard when it comes to vocabulary production (Johnson E. , 2015). Many classroom standards are set in accordance with middle- and upper-class SES language norms in mind (Johnson E. , 2015). This makes it extremely difficult for lower income students from diverse backgrounds to break through these academic hurtles (Johnson E. , 2015). Johnson disagreed with the socialization mismatch hypothesis which posits that children have a better chance to succeed at school when their home language and reading abilities match those accepted in the educational environment (Johnson E. , 2015). Johnson stressed that low-income students may have a hard time in academics because of unfair influences of more affluent cultural groups, not because they are linguistically or mentally disadvantaged (Johnson E. , 2015).

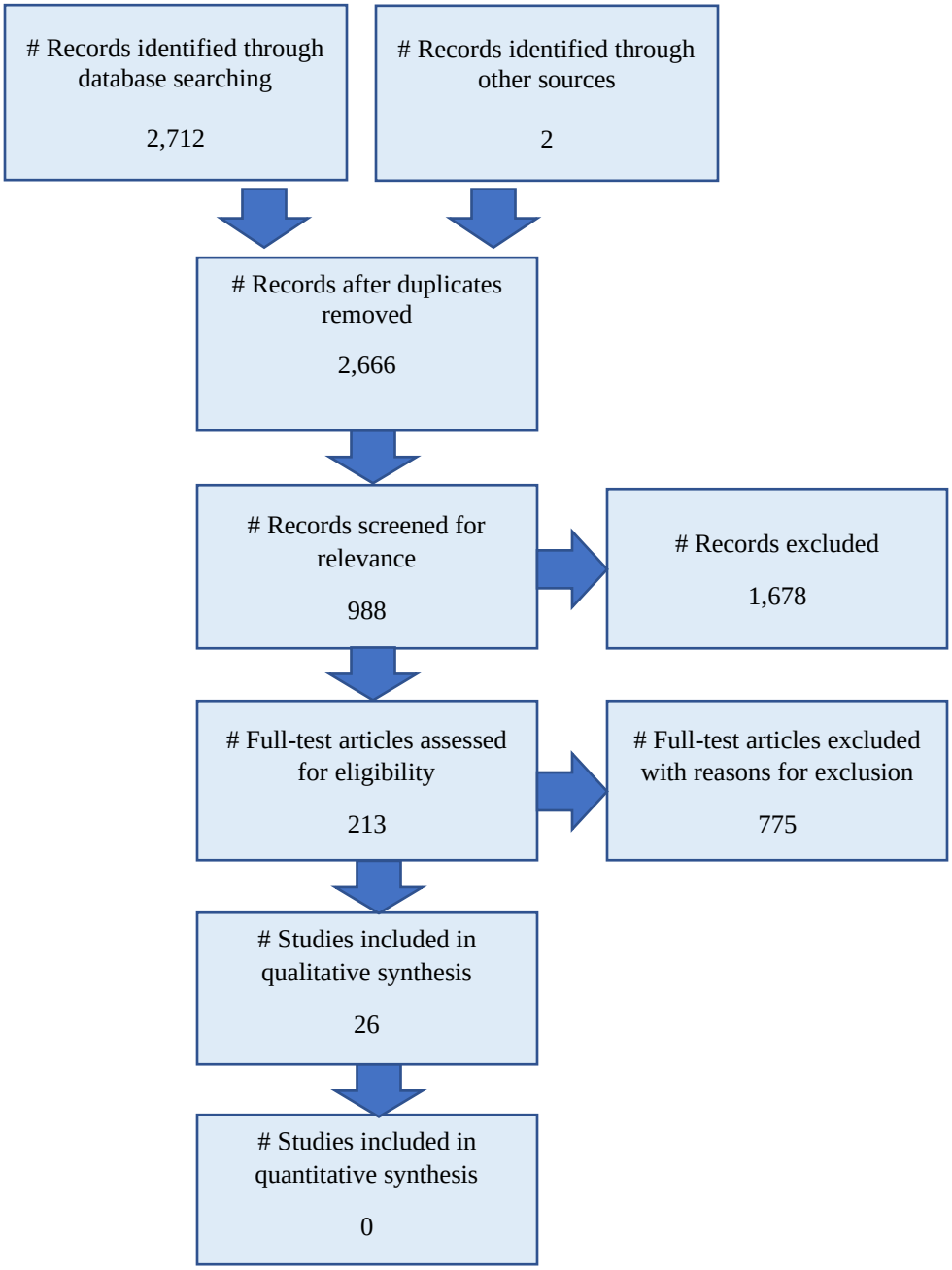
Sperry (2018), and colleagues attempted to debunk the claim that low-income children hear 30 million less words than middle-income children during the beginning stages of life. Their five longitudinal studies included data on ethnographic information in conjunction with home observations of 42 children from 18-48 months of age (Sperry, 2018). The connection between words spoken by the primary caregiver to the child showed no discernable link in regard to socioeconomic status (Sperry, 2018). In particular, the children in the low-income Black Belt community heard more words spoken than most children in middle-income communities (Sperry, 2018). The study revealed that in middle-income, working class, and poor communities, the amount of words spoken to the child by the primary caregiver and words spoken by all

caregivers increased (Sperry, 2018). The different social classes did not influence how much language a child was exposed to (Sperry, 2018). Finally, when the combined speech a child heard was analyzed the results showed children in poor communities heard 3,203 words per hour (Sperry, 2018). This was more words than the middle-income group of children heard and showed once again that linking low socioeconomic status and language input is a dubious proposition at best (Sperry, 2018).

III. METHODOLOGY

The method used was taken from the Preferred Reporting Items for Systematic Reviews the PRISMA method.

Figure 1: PRISMA selection process



3.1 Protocol and registration

A current review protocol registration for this study is not registered

3.2 Eligibility criteria

The eligibility criteria for this study was organized using the following criteria. Population: Spanish monolingual, English monolingual, and Spanish and English Bilingual toddlers from 2-36 months of age. Intervention used: The MacArthur-Bates Communicative Inventory English and Spanish versions. Factors effecting vocabulary: Socioeconomic status, ethnicity, parental education and disability. Also, any issues with format or design of CDI were included. The studies were from 1980-2019. All 26 studies reviewed have been published in academic journals and are within the 1st and 2nd quartile for research credibility.

Table 1: Eligibility criteria

Author	Population	Intervention	Objective
(Arriaga, Fenson, Cronan, & Pethick, 1998)	103 English speaking children 16-30 months	MacArthur-Bates CDI (English)	Compared the language skills in a group of very low-income toddlers with those of a middle-income sample
(Bates, 1994)	1,803 English speaking infants aged between 0; 8 and 2; 6	MacArthur-Bates CDI (English)	Developmental aspects of vocabulary composition. Looked at effects of parental education, SES
(Brady & Goodman, 2014)	48 English speaking toddlers 18-36 months	MacArthur-Bates CDI (English)	Vocabulary production, maternal education looked at
(Checa, 2016)	108 Down Syndrome children compared to 108 typically developing children that speak Spanish	MacArthur-Bates CDI (Spanish)	Vocabulary production between DS and TD compared. Maternal education looked at
(Core, 2013)	47 Spanish-English bilingual toddlers 22-30 months	MacArthur-Bates CDI (Spanish and English)	Vocabulary growth both total and conceptual. CDI design

(de Diego-Lazaro, 2019)	53 Spanish toddlers 8-34 months who are deaf or hard of hearing	MacArthur-Bates CDI (Spanish)	Identify predictors of expressive vocabulary in young Spanish-speaking children who are deaf or hard of hearing living in the United States. Looked at maternal education
(Feldman, et al., 2000)	2,156 monolingual English-speaking toddlers 1-2 years old	MacArthur-Bates CDI (English)	Vocabulary production with effects of maternal education, race
(Foster-Cohen, J, Champion, & Woodward, 2007)	90 English speaking pre-term toddlers and 102 full term toddlers at 2 years old	MacArthur-Bates CDI (English)	vocabulary production with effects of SES and maternal education
(Furey, 2011)	23 English speaking toddlers 16-18 months	MacArthur-Bates CDI (English)	Vocabulary production between low and middle-income families
(Gale, 2008)	15 English speaking toddlers, 2 years of age. 5 deaf toddlers with deaf parents, 5 deaf with hearing parents, and 5 hearing with hearing parents	MacArthur-Bates CDI (English)	vocabulary production in deaf and hard of hearing measured with CDI
(Gulberson, 2011)	45 Spanish speaking toddlers, 2 years of age	MacArthur-Bates CDI (Spanish)	Effects of maternal education on vocabulary production
(Heilmann, 2005)	138 English late talking and normally developing toddlers 30 months	MacArthur-Bates CDI (English)	Validity of CDI. Looked at maternal education effects
(Houston-Price, 2007)	30 English speaking toddlers with a mean age of 1 year and 6 months	MacArthur-Bates CDI (English)	Validity of Parents scoring of CDI. Issues with understanding instructions
(Hurtado, 2014)	37 bilingual English and Spanish speaking toddlers at 30 months	MacArthur-Bates CDI (English and Spanish)	Dual language toddler's performance on the CDI. SES and education factors looked at
(Jackson-Maldonado. D., 2013)	Spanish speaking toddlers, 718 from middle class families and 1818 children from low income families	MacArthur-Bates CDI (Spanish)	Comparing vocabularies of middle- and lower-class children. Effects of SES and maternal education
(Luyster, 2007)	93 English speaking toddlers with autism, 31 with developmental delay, and 29 typically developing. 2 years old	MacArthur-Bates CDI (English)	Validity of CDI with analyzing vocabulary in children with autism and developmental delay

(Mancilla-Martinez, 2011)	79 bilingual Spanish and English toddlers, 24-36 months	MacArthur-Bates CDI (Spanish and English)	Utility and validity of the MacArthur-Bates Communicative Development Inventory (CDI) for use with low-income parents and their 24- to 36-month-old Spanish-English bilingual children
(Marchman, 2002)	28 bilingual English and Spanish toddlers, 8-34 months	MacArthur-Bates CDI (Spanish and English)	Validity of two analogous caregiver/parent report measures of early language development in young children who are learning both English and Spanish. Effects of maternal education
(McDuffie, 2005)	29 English speaking 2 and 3-years with autism	MacArthur-Bates CDI (English)	Predict vocabulary production in children with autism
(Pan & Rowe, 2004)	105 English speaking toddlers 2 years	MacArthur-Bates CDI (English)	Low-income families scored their monolingual English children at 2 and 3 months of age on the MacArthur-Bates CDI. Also looked at effects of race and location of home
(Pearson, 1994)	20 bilingual English and Spanish toddlers, 10-30 months	MacArthur-Bates CDI (Spanish and English)	Bilingual vocabulary compared to monolingual production
(Ramírez, 2019)	71 English speaking toddlers, 6-18 months	MacArthur-Bates CDI (English)	Effects of parent coaching and vocab abilities. SES effects included
(Roberts, 1999)	87 English speaking toddlers, 18-30 months	MacArthur-Bates CDI (English)	Vocabulary scores on the MacArthur-Bates CDI of monolingual English speaking African American children from low-income families
(Scherer, 1999)	3 English speaking toddlers with clef-palate	MacArthur-Bates CDI (English)	Vocabulary scores on CDI effected by disability
(Thal, 1999)	12 English speaking children, 24-32 months	MacArthur Bates CDI (English)	Validity of parent report with toddlers with language delay
(Yoder, 1997)	17 English speaking toddlers mean age of 25 months with developmental delays	MacArthur-Bates CDI (English)	Vocabulary production measured by CDI. Effects of maternal education

3.3 Information sources

The databases used for this synthesis were, ERIC, Psych INFO, Google Scholar, and Web of Science. ERIC is an online library of education research and information. It is accredited and sponsored by the U.S. Department of Education and the Institute of Education Sciences. All search data from ERIC was looked up from the week of February 10th-Feburary 15th, 2020. Psych INFO is a database that provides access to peer-reviewed articles in the realm of behavioral and social sciences. All search data from Psych Info was looked up from the week of February 10th-February 15th, 2020. Google Scholar is a database the provides access to peer-reviewed articles and research studies. All search data from Google scholar was looked up from November 10th, 2019 to February 20th, 2020. Web of Science is a database that offers information on various educational areas. It was created by the Institute for Scientific Information and is run by Clarivate Analytics. All search data from Web of Science was looked up from the week of February 10th-February 20th, 2020.

3.4 Search Strategy

The research for this study consisted of peer-reviewed articles taken from databases including, ERIC, Psych INFO, Google Scholar, and Web of Science. Factors that were addressed included SES, issues with design or instructions of assessment, and education levels of parents involved. Here is an example of search parameters used in the Psych INFO database: Search terms included terms (infant* OR toddler* OR child*) AND (2-36 months of age* OR pre-school*) AND (validity* OR effectiveness* OR utility*) AND (socioeconomic status* OR maternal education* OR parental education* OR low-income* OR middle-income* OR high income*) AND (English monolingual* OR Spanish monolingual* OR (Spanish and English bilingual* AND (typically developing* OR late-talk* OR developmental disability* OR

language delay*) AND (parent-based* OR home* OR home-based*) The option peer-review was selected for all search databases.

3.5 Study selection

The current synthesis had certain parameters that were considered to narrow down the analysis. First, the research reviewed only involved English monolingual, Spanish monolingual or English and Spanish bilingual speaking children. Any research that involved other languages besides these were omitted. The data is from the years 1980-2019. Second, studies that involved children who had been diagnosed with preexisting conditions such as Down Syndrome, Autism, deaf, hard of hearing or clef palate were separated into their own section. Third, the MacArthur-Bates CDI needed to be used as a part of the research. Any studies where the participants only used this test to qualify for the study were omitted. Finally, studies were refined to include infants and young children from 2 months to 36 months of age.

3.6 Data collection process

For this synthesis all data was extracted independently by the researcher. All data was collected via research databases available through access through Florida International Universities research library database. The authors of the studies were not contacted for any results-based purposes. All results were extracted directly from the studies with proper citation and credit given to the authors.

3.7 Data items

Data was looked at for the following variables: socioeconomic status, parental education, race, design issues, disability factors. For the purposes of this study, socioeconomic status is defined at the income level of the family of the participant in the study. Parental education describes the level of education of the parents of the participants in the study. Race includes the

race of the parents and participants in the study. Design issues include any problems with the assessment brought up by either the parents of the participants or the researchers. These may include issues with instructions or format of assessment. Disability includes any physical or mental disabilities that the participants have while participating in the studies.

3.8 Risk of bias in individual studies

The studies were checked on their ability to apply appropriate eligibility criteria, valid measurements and outcomes, and their ability to control any confounding variables. An example of eligibility criteria for a study would be to include Spanish or English speaking children between 2-36 months and to include a variable like socioeconomic status, race, design, disability, or parental education. The study also had to use the MacArthur-Bates CDI as a primary part of the study. Arriaga, Fenson, Cronan, and Pethick (1998), met these criteria because they dealt with English speaking children from 16-30 months and looked at the effects of SES on the results of the MacArthur-Bates CDI. It also used the MacArthur-Bates CDI as a primary part of the experiment (Arriaga, Fenson, Cronan, & Pethick, 1998). Valid measurements and outcomes were checked by looking at face validity and predictive validity. In Arriaga, Fenson, Cronan, and Pethick (1998), the face validity was good because the parents believed that the CDI was testing what it was intended to access. The predictive validity was not valid because SES effected testing results (Arriaga, Fenson, Cronan, & Pethick, 1998). This meant that making any valid predictions on child vocabulary acquisition or future language delays would be compromised. Studies controlled for confounding variables by using tests like ANOVA and using stratified sampling techniques to include valid test participants.

3.9 Summary of measures

Studies used various measures to describe their data including measures of central tendency like mean, and mode. For example, Foster-Cohen, J, Champion, and Woodward (2007), compared the means of vocabulary production between very preterm infants and non-preterm infants. Studies also used measures of variation including range, standard deviation, and variance. Marchman (2002), and colleagues looked at the range of vocabulary scores between children learning Spanish and English at 26 months. Yoder (1997), looked at the standard deviations of item by item validity when showing the amount of total words understood.

IV. RESULTS

For this study there were initially 2,712 studies that came up on the initial search for the MacArthur-Bates CDI. These studies were identified through academic searches through various research databases through the Florida International University Library website. Next the studies were filtered into only those that were peer-reviewed and in the 1st or 2nd quartile of article strength. Duplicates were removed as well. The titles of the studies and abstracts were reviewed, and 213 studies remained. Then the studies were matched against the search criteria that was provided. Studies were removed for several reasons: Study did not use MacArthur-Bates CDI English or Spanish as a primary measure, Study looked at any language other than English or Spanish, Study did not contain information about effect of SES, parental education, race, design, or disability on MacArthur-Bates CDI. Study did not fall within the proper date range, study contained children that were not in the proper age range, 24 studies remained after this. Two additional studies were added later to bring the total to 26 studies.

4.1 Intervention description

The following studies were sorted into three areas of emphasis. The three areas were, participants, setting, and length of study.

4.1.1 Intervention emphasis

The three areas emphasized in the interventions were, vocabulary development, parent report accuracy, and parent coaching effectiveness.

4.1.2 Vocabulary development

Most of the studies that were analyzed dealt with how much vocabulary the participants could produce. (88%, n=23) of the studies fell into this intervention emphasis. This was done

using the MacArthur-Bates Communicative Developmental Inventory. For example, Scherer (1999), used the MacArthur-Bates CDI to measure the vocabulary production of three 2-year old monolingual English speakers with developmental delays.

4.1.3 Parent Report Accuracy

A few of the studies dealt with parent report accuracy. (0.08%, n=2) of the studies fell into this category of intervention emphasis. For example, Marchman (2002), and colleagues looked at the utility of the MacArthur-Bates CDI with 26 Spanish and English bilingual toddlers from 24-36 months of age.

4.1.4 Parent Coaching Effectiveness

Only one of the studies emphasized parent coaching effectiveness. A study by Ramirez (2019), looked at whether parent coaching at 6, 10, and 14 months can affect parental language input which in turn can affect language development.

4.1.5 Risk of bias within studies

The studies in this synthesis were assessed for risk by checking their face validity and predictive validity. Face validity refers to whether a test looks like it is measuring what it is intended to measure (Mcleod, 2013). Predictive validity is how well the scores on a test can accurately forecast what the test is trying to show (Ryan, 2015). These two types were chosen for a couple of reasons. It was important to know how confident the parents felt about using the MacArthur-Bates CDI and it was important to see if the CDI could do its intended purpose. Its purpose is to help predict future language delay. The results show that face validity was high in almost all the studies accept two. Most parents believed they understood the test and that it was carrying out its intended purpose. Twelve of the studies showed interference from other variables like SES, parental education, disability, race, and design affected how well vocabulary scores

were assessed. If the vocabulary skills are compromised, prediction of language delay is compromised.

Table 2: Risk of bias within studies

Study	Appropriate eligibility criteria	Valid measurements and outcomes	Control for confounding variables
(Arriaga, Fenson, Cronan, & Pethick, 1998)	Population: English. 16-30 months Relevant variables: SES MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, Correlation between SES and results	ANOVA Random sampling
(Bates, 1994)	Population: English. 8-30 months Relevant variable: Education, SES MacArthur-Bates primary assessment tool: Yes	Face Validity: No Predictive Validity: No, SES affected results	ANOVA Stratified sampling
(Brady & Goodman, 2014)	Population: English. 18-36 months Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, correlation between education and results	ANOVA Stratified sampling
(Checa, 2016)	Population: Spanish: >3 years Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes Disability did not affect results	ANOVA Stratified sampling
(Core, 2013)	Population: Bilingual: 22-30 months Relevant variable: Design MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, bilingual vocabulary lower than monolingual	ANOVA Stratified sampling
(de Diego-Lazaro, 2019)	Population: Spanish: 8-34 months Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes. No correlation between education and results	Stratified sampling Hierarchical multiple regression analysis

(Feldman, et al., 2000)	Population: English: 1-2 years Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, education had correlation with results	ANOVA Stratified sample
(Foster-Cohen, J, Champion, & Woodward, 2007)	Population: English: 2 years Relevant variable: SES MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	ANCOVA Stratified sample
(Furey, 2011)	Population: English: 16-18 months Relevant variable: SES MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, SES did not affect results	ANOVA Cohen's guidelines Stratified sample
(Gale, 2008)	Population: English: 2 years Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	ANOVA Stratified sampling
(Gulberson, 2011)	Population: Spanish: 2 years Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	Pearson product-moment correlations t-test Stratified sample
(Heilmann, 2005)	Population: English: 30 months Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, education did not affect results	Stratified sample One tailed Pearson correlation coefficient False detection rate method
(Houston-Price, 2007)	Population: English: 1 year, 6 months Relevant variable: Design MacArthur-Bates primary assessment tool: Yes	Face Validity: No Predictive Validity: No, design affected results	ANOVA Stratified sample
(Hurtado, 2014)	Population: Bilingual: 30 months Relevant variable: SES, Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, education and SES did not affect results	Stratified sampling

(Jackson-Maldonado. D., 2013)	Population: Spanish: >3 years Relevant variable: SES, Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, SES affected results	ANOVA Stratified sampling
(Luyster, 2007)	Population: English, 2 years Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	ANOVA Stratified sampling
(Mancilla-Martinez, 2011)	Population: Bilingual: 24-36 months Relevant variable: SES MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, SES did not affect results	Stratified sampling
(Marchman, 2002)	Population: Bilingual: 8-34 months Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, education did not affect results	Stratified sampling
(McDuffie, 2005)	Population: English: 2-3 years Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	Matched sampling Principle component analysis Multiple regression analysis
(Pan & Rowe, 2004)	Population: English: 2 years Relevant variable: SES, Race MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, race did affect results	ANCOVA Stratified sampling
(Pearson, 1994)	Population: Bilingual: 10-30 months Relevant variable: Design MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, design affected results	Stratified sampling
(Ramírez, 2019)	Population: English: 6-18 months Relevant variable: SES MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, SES did not affect results	Hollingshead Index scores ANOVA Stratified sampling

(Roberts, 1999)	Population: English: 18-30 months Relevant variable: SES (Core, 2013) MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, education affected results	ANOVA Stratified sample
(Scherer, 1999)	Population: English: 2 years Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: Yes, disability did not affect results	Matched sample Percentage agreement strategy
(Thal, 1999)	Population: English: 24-36 months Relevant variable: Disability MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, disability affected results	Stratified sample Statview used to organize data
(Yoder, 1997)	Population: English: Education Relevant variable: Education MacArthur-Bates primary assessment tool: Yes	Face Validity: Yes Predictive Validity: No, education affected results	Cohen's kappa Stratified sample

4.1.6 Results of individual studies

The following two tables have three areas of focus. The first is whether the MacArthur-Bates CDI proved effective in its ability to accurately access the vocabulary skills of the participants of the study. The second area shows the outcome of each of the studies and main points made by the researchers. It also represents a written description of the statistical data. The final area is the statistical data that was described in the results section. The first table shows only the studies performed on participants that had a disability. The second table shows results of participants with no disability.

Table 3: Results of individual studies with disability

CDI with Disability	CDI effectiveness	Results	Statistics
(Checa, 2016)	Yes,	CDI provides info about words known, but not about their frequency of use. Children may produce words but use them infrequently. No significant influence of maternal education Down Syndrome produced more nouns than typically developing children	Down Syndrome (Mean nouns = 39.21) Typically Developing (Mean nouns= 34.40)
(de Diego-Lazaro, 2019)	Yes	Maternal Education did not significantly correlate with vocabulary production Younger children scored higher than older on vocabulary	Maternal education r (53) = .13, p > .001 positive effect of early age of intervention on vocabulary quotients decreased as chronological age increased F (5, 45) = 16.94, p < .001
(Foster-Cohen, J, Champion, & Woodward, 2007)	Yes	The CDI did accurately show that children born very preterm are at a elevated risk for language delays in their language development The study showed no between group differences in socioeconomic status or maternal education when it came to vocabulary scores	Extremely Preterm= Mean=181.22 Very Preterm= Mean= 250.13 Full Term= Mean = 260.89
(Gale, 2008)	Yes	Deaf toddler produced significantly less words than hearing toddlers	Hearing toddlers = mean of 128.4 words Deaf toddlers = mean of 53.4 words
(Gulberson, 2011)	Yes	There was strong validity for the CDI with Spanish speaking toddlers with down syndrome from 24 to 35 months of age	For two CDI-based measures, 1/3 of children who are classified as delayed by the screening measure will have this confirmed by the SPLS–4

(Luyster, 2007)	Yes	The autism group scored lower than the two other groups on verbal IQ, nonverbal IQ, words known, and phrases understood	Verbal: autism: *Mean*=30·88, *Standard Deviation*=23·32 PDD-NOS: *Mean*=60·52, *Standard Deviation*=30·81; *t* (94·14) =-5·50, *p*<0·001 non-verbal IQ (autism: *Mean*=58·52, *Standard Deviation*=23·38 PDD-NOS: *Mean*=78·30, *Standard Deviation*=22·89; *t* (92·95) =-4·25, *p*<0·001
(McDuffie, 2005)	Yes	None of the participants reached the ceiling level in number of words understood or number of words understood and said, indicating that the CDI Infant Scale was developmentally appropriate for this group of children	Variable Mean SD Range Time 1 CDI comprehension 125.93 96.93 6-372 Time 1 CDI production 61.83 95.18 0-346 Time 2 CDI comprehension 225.79 102.91 0-390 Time 2 CDI production 137.62 122.40 0-390
(Scherer, 1999)	Yes	The CDI showed improvement pre-test and post-test on all three areas of language measured including the number of vocabulary words, mean of the three longest sentences, and number of suffixes used	Vocabulary went from 5th percentile to 30th from pre-test to post test

(Thal, 1999)*	Issues with comprehension and gesture scores	CDI was applicable to assessing vocabulary in young toddlers with language delay, however, comprehension scores and gestures were not as valid	Number of different words produced .66[*] .52[**] - .30 - .27 -.41 .41
(Yoder, 1997)	Issues with what words a child understands	CDI shows adequate stability dealing with total number of words which helps with identifying vocabulary delay Clinicians should not relay on results from what words a child understands as SES from low-income parents may affect results	Reliability only .60 on nouns, games and action words Normally .80-.90 reliability is range for an accurate clinical decision to be made

*For Thal (1999), a special words and gestures form of the CDI was used. This includes two parts, 1) parent records words child understand and produces. 2) Parents record any symbolic gestures child has attempted

Table 4: Results of individual studies without disability

CDI without Disability English	CDI Effectiveness	Results	Statistics
(Arriaga, Fenson, Cronan, & Pethick, 1998)	Issues with SES effecting CDI results	There was small negative relationship between the vocabulary production and socioeconomic status	Low income Middle income Mean Mean 212.50 359.31 210.07 276.37
(Bates, 1994)	Issues with education and SES Issues with design	Small correlation in favor of higher education and SES producing higher vocabulary parents had difficulty differentiating between their child understanding a word and just being able to repeat it	maternal education (r = +O' 1 I, P < 0'05) paternal education (r = +O'II, P < 0'05) paternal occupation (r = +0'10, P < 0'05).

(Brady & Goodman, 2014)	Issues with maternal education effecting scores	Results showed that the higher the maternal education the lower percentile the child scored in on all four stages of testing on vocabulary	Maternal Education Years 16 16.3 16.6 16.8 Months 18 24 30 36 60.0 59.7 48.2 40.7
(Feldman, et al., 2000)	Issues with education effecting CDI results	Mothers with lower education recorded higher scores for their children than mothers with high education for vocabulary	Education <High School= 115.4 High School= 100.9 College= 91.0
(Furey, 2011)	Yes	No difference in accuracy in maternal reporting for low-income and middle-income children at both 16 and 18 months	Low income 16 months 18 months 120 135 Middle income 174 282
(Heilmann, 2005)	Yes	Data was in favor of the utility of the CDI in predicting language out outcomes in late talking toddlers	24 months late talking toddlers had scores below the 10th percentile on productive vocabulary, however scores rose to above the 15th percentile at 30 months
(Houston-Price, 2007)	Issues with parent reporting, Design	Researchers discovered that parents reported their children comprehended many unknown words and suspected that they may be overestimating their vocabulary abilities studies were rife with false negatives in infant vocabulary production as the rate of comprehension of known words and unknown words was relatively the same	Hearing both known ($t(28) = 2.46$, $p = 0.02$) and unknown words ($t(28) = 2.46$, $p = 0.02$).
(Pan & Rowe, 2004)	Issues with ethnicity and community	Race and ethnicity did influence CDI scores with white mothers scoring their 2-month old's higher than black and Hispanic mother. Rural children scored higher than urban children on the CDI	Black ($p<0.05$) and Hispanic mothers ($p<0.01$) Rural vs. Urban CDI ($F(2,102) = 8.52$, $p<0.01$

(Roberts, 1999)	Issues with home environment effecting CDI results Underreporting issues	Children that came from home environments that are more supportive had higher vocabulary scores, regular nouns and verb scores, and longer utterances	Scores for nouns and verbs dropped from the 68th percentile at 18 months to the 55th percentile at 24 months and continued to drop to the 31st percentile at 30 months
CDI without Disability Spanish	CDI Effectiveness	Results	Statistics
(Jackson-Maldonado. D., 2013)	Issues with SES effecting CDI results	In the younger age group, children from low SES backgrounds were assessed as having a higher number of words understood than children with higher SES Vocabulary comprehension showed lower scores for children in the younger age group that had educated families and higher in younger children from uneducated families	F (1, 593) = 10.7, p < .005
CDI without Disability Bilingual	CDI Effectiveness	Results	Statistics
(Core, 2013)	Issue with design	Taken separately, Spanish and English bilinguals conceptual vocabulary was lower than monolingual English and Spanish speakers. Total vocabulary of bilinguals was within normal levels of production when compared to monolingual English and Spanish speakers	At 22 months, conceptual vocab of bilingual children fell below 25th percentile At 22 months, total vocab of bilingual children at or below 25th percentile matched monolingual children at that mark

(Hurtado, 2014)	Yes	Expressive and receptive vocabulary scores recorded showed toddlers knew more Spanish than English words at 30 and 36 months	30 months, r (36) = .59, p < .001 36 months, r (29) = .62, p < .001
(Mancilla-Martinez, 2011)	Yes	Low income parents could distinguish between words their child said and words they understood	Words understood means fell within norms
(Marchman, 2002)	Yes	Factors like maternal education, what language was spoken at home. proportion of English to Spanish output and mother's acculturation level did not influence CDI vocabulary scores	Mother's Years of Education (English: r = .66, Spanish: r = .82) Proportion of English-to-Spanish input (English: r = .69, Spanish: r = .73) Mother's Acculturation Level (English: r = .53, Spanish: r = .76)
(Pearson, 1994)	Issues with design	Recording scores of Spanish and English bilingual toddlers separately put them below typically developing monolingual toddlers in terms of vocabulary scores Total vocabulary scores were more within normal developmental range	At 27 months separate scores: Spanish: 160 English: 80 27 months total combined score: Spanish/English: 240
(Ramirez, 2009)	Yes	No correlation between SES and CDI performance Parent coaching did help performance	F (2.74, 183.40) = 2.96, P = 0.038, F (1.86, 122.53) = 4.69, P =0.0130.07; F (2.00,131.93)=6.22,P=0.003..

4.2. Participants

This section talks about all the individuals that are given intervention in these studies. This includes the parents and toddlers that took part in the studies. This includes information on the following if it was given, design, socioeconomic status, education, race, and disability status.

4.2.1 Socioeconomic status

Most of the studies looked at some form of Socioeconomic status when conducting their research. (69%, n=18) looked at low- or middle-income families to see the effects this had on CDI results. The socioeconomic status of mothers was recorded when observing how toddlers with autism performed on the CDI (McDuffie, 2005). Five studies brought up issues with socioeconomic status effecting CDI results in a negative way. Jackson-Maldonado, D (2013), and colleagues pointed out unusual correlations about socioeconomic status effecting results. Certain families with young children in low SES households were scoring their children higher on vocabulary than children in families with middle class SES.

4.2.2 Education

A few of the studies reviewed looked at some level of either parents' education. (42%, n=11) of the studies looked at parental education and its influence on the CDI results. The link between parental education achievement and other variables was observed in conjunction with CDI toddler performance (Ramírez, 2019). Only five of these studies provided a negative correlation between parent's education level and negative performances in vocabulary production on the CDI.

4.2.3 Race and Ethnicity

Less than half of the studies did not disclose the race of the participants (46%, n=12). (.125%, n=3) had reported white ethnicity, (.125%, n=3) reported Euro-American, (.125%, n=3)

reported Asian, (25%, n=6) reported African American, (33%, n=8) reported Latino, and (.125%, n=3) reported other, or multiracial. Only one study by Pan, highlighted race as having a effect on CDI results. Pan and Rowe (2004), showed evidence that white mothers were scoring their 2-year-old vocabulary scores higher than African American and Hispanic mothers.

4.2.4 Disability

Most of the studies did not include toddlers with disability (61%, n=16). (40%, n=10) of the studies reported on toddlers with disability. two studies reported on autism, two reported on hard of hearing or deaf toddlers, and two reported on toddlers with language delays. Other disabilities included, clef-palate (Scherer, 1999), pre-term babies (Foster-Cohen, J, Champion, & Woodward, 2007), and down syndrome (Checa, 2016). Most of the studies showed that the MacArthur Bates CDI was a valid tool to use when assessing a child 2-36 months with a learning disability.

4.2.5 Design

Only 3 of the studies (.11 %, =3) talked about how design could be improved on the CDI to help with better reporting and results. Issues included how the CDI should encourage bilingual parents to count total vocabulary and how instructions can be confusing as far as what it means to have a child understand a word vs just be able to produce it.

4.2.6 Intervention setting

All the studies reviewed had the MacArthur-Bates CDI reported on by parents at the home of the child. Instructions on how to complete the CDI were either given to the parents over the phone mailed or explained in person. For example, in Bates (1994), the parents were asked to complete the CDI at home and mail in the results. The results were analyzed by a clinician and then fed through a machine that recorded and checked for any errors the parents may have made

(Bates, 1994). In Marchman (2002), the parents were mailed their CDI forms and completed them at home. Then they turned them in during their visits to the lab (Marchman, 2002). In this instance the parents received instructions over the phone on how to complete the CDI while also being mailed a paper version of separate instructions in both English and Spanish (Marchman, 2002).

4.2.7 Length of Intervention

The length of the time it took to complete the MacArthur-Bates CDI's for the studies varied between one session lasting the course of a day to multiple sessions lasting 16 months. The study by Bates (1994), involved the parents completing the CDI at home for one session and turning it in to be evaluated. In the study conducted by Brady and Goodman (2014), the mothers were asked to report their child's scores on the CDI at intervals that lasted the course of 16 months.

4.2.8 Languages used in studies

The languages of the children that were focused on in this synthesis were English monolinguals, Spanish monolinguals, and English and Spanish Bilinguals. Of the 26 studies reviewed (61%, n=16) of the studies had English monolingual speaking children. (15%, n=4) had Spanish monolingual speaking children and (19%, n=5) had English and Spanish Bilingual speaking children. For example, Marchman (2002), used the MacArthur-Bates CDI to access 26 Spanish and English bilingual toddlers from 24-36 months of age.

4.3 Measures

All the studies used the MacArthur-Bates CDI to access potential language delays in children. All studies used either the English or Spanish versions of the CDI. One study by Pan and Rowe (2004), also used the Peabody Picture Vocabulary test as well as the MacArthur-Bates

CDI to compare and contrast results. (61%, n=16) of studies used the English only version of the MacArthur-Bates CDI to test language skills. (19%, n=5) studies used the monolingual Spanish version of the MacArthur-Bates CDI to access child language skills. (19%, n=5) of studies used Spanish and English forms of the MacArthur-Bates CDI for language assessment.

4.4 Intervention effectiveness

This section will describe the effect the MacArthur-Bates CDI interventions had on accurately accessing child language skills in the studies. Studies looked at how outcomes of vocabulary production including known and unknown words, expressive vocabulary, receptive vocabulary, and mean length of utterance were impacted by intervention. (46%, n=12) of the studies reported that the MacArthur-Bates CDI provided effective results when it came to language outcomes. (53%, n=14) of the studies found issues with the validity of the results on the CDI with regards to several factors including SES, design of the assessment, parent underreporting, and ethnicity of the families reporting the results.

4.4.1 Effectiveness on English outcomes

Of the 16 studies that were done using the English version of the MacArthur-Bates CDI, (28%, n=7) were found to yield results that were considered valid. The remaining studies reported issues with socioeconomic status, design of CDI, maternal or parental education, ethnicity, and parent under reporting. For example, in Yoder (1997), the researchers cautioned clinicians on making decisions on whether a child may need therapy because results on words a child understands was compromised by low-income maternal reports. The validity of scores by the low-income mothers was .60 which is low for an accurate clinical decision to be made (Yoder, 1997). In Pan and Rowe (2004), ethnicity played a role in the outcomes of CDI results. White mothers scored their 2-month old's higher than black and Hispanic mothers.

4.4.2 Effectiveness on Spanish outcomes

Of the four Spanish version MacArthur-Bates only two were found to be valid. The other studies had issues with design of the CDI, and socioeconomic status. For example, in Checa (2016), the researchers noted that the CDI provides info about words known, but not about their frequency of use. Children may produce words, but use them infrequently so the results may not be an accurate representation of their real vocabulary repertoire (Checa, 2016).

4.4.3 Effectiveness on Bilingual outcomes

Of the six Spanish and English bilingual studies observed, only issues with design were brought up regarding effecting accurate results. Pearson (1994), mentioned that the CDI's only record vocabulary scores of Spanish and English children separately which may put them below the language norms of monolingual speakers. A total vocabulary score should be counted to gain an accurate read on how bilingual children are progressing in comparison to their monolingual peers (Pearson, 1994).

V. DISCUSSION

The purpose of this synthesis was to systematically review studies that had used the MacArthur-Bates CDI to access the language skills of children from 2-36 months of age. The synthesis focused on Spanish and English monolingual and Spanish and English bilingual toddlers. This population has not yet been reviewed in a synthesis to the knowledge of this researcher. The aims of this study were to see what factors might affect the results of MacArthur-Bates CDI and to determine if this parent-based assessment can be relied upon for predicting language delay in a toddler.

The results in just under half of the studies analyzed brought up some issue with the accuracy of the scores obtained. This synthesis has demonstrated that outside variables and even certain design aspects of the MacArthur-Bates CDI can compromise the data obtained. However, many studies admitted that any parent-based assessment must be interpreted with a bit of caution. There were some strengths to the CDI that were discovered during this synthesis including its effectiveness in interpreting vocabulary results in toddlers with disabilities, the overall flexibility the CDI provides as far as how it can be administered and the consistent settings in which the CDI was administered.

5.1 Summary of the current body of research

This study identified two main strengths of the MacArthur-Bates CDI that showed itself in the current research. Its effectiveness in accessing vocabulary skills in toddlers with disability and the advantages in how it can be administered. The MacArthur-Bates CDI has a few pros to how it can be administered compared to other assessments. One is the fact that it can be used longitudinally (Furey, 2011). Also, inconsistent application of the test does not interfere with the

results (Furey, 2011). The fact that it can be administered by parents allows for a significant amount of statistics to be collected (Arriaga, Fenson, Cronan, & Pethick, 1998). More data that's available for analysis leads to greater validity of studies (Arriaga, Fenson, Cronan, & Pethick, 1998).

The CDI also offers advantages when it comes to some of the issues that can arise when young children are given assessments (Heilmann, 2005). These can include the fact that toddlers don't always communicate a lot and therefore make it difficult to acquire language samples (Heilmann, 2005). It also allows for the child to be assessed by someone they know which can elicit more willingness to corporate (Heilmann, 2005). Another strength of the study is that all the settings for which the CDI was administered was in the home. This creates consistency of across all the studies in terms of the test environment.

The CDI showed considerable effectiveness when interpreting vocabulary skills in toddlers with a disability. Of the ten studies that involved a toddler with a disability, seven showed accurate data. In most studies researchers were able to get higher vocabulary scores when comparing toddlers with disability and those without. In fact, in the three studies that found issues, two of the studies had more of a contention with the design of the CDI itself and not the results acquired.

5.1.1 Socioeconomic status

Various studies attempted to show how the MacArthur-Bates CDI was affected by socioeconomic status. Arguments were made that over or underreporting may have occurred because families may have misinterpreted results or were too harsh on their child's language skills. The possibility that lower-income families had less language output was also offered as an excuse for lower vocabulary scores. It is important to empathize caution when reviewing these

opinions. The link between SES and language performance has been shown to be a weak one. As stated earlier in the review of literature, studies have proven that low SES children from low SES families experience just as much if not more language input than middle SES families. The reasons for higher vocabularies or lower vocabularies based on SES in these studies are purely speculation by the researchers and should not be accepted as true.

5.1.2 Education

Some correlations between higher educated parents scoring their child lower than less educated parents caused researchers to question whether higher educated parents had higher standards for their children causing them to be stricter during assessments. At the same time parents with less education scored their children lower than those with higher education in some studies. Education levels did play a part in effecting some CDI results, but once again any reasons given by the researchers are speculation at best. It is clear that more research needs to be done to understand why there is such an unpredictable dynamic when it comes to parental education and CDI results.

5.1.3 Race

Race was not a big factor when determining results on CDI performance. Only one study focused on race and did not offer any valid reasons for why it effected CDI results.

5.1.4 Design

Design of the CDI was brought up as effecting results on a few occasions. The main idea being that for bilingual children the CDI should be structured to report on total vocabulary to provide families with a more accurate score. Only looking at separate scores on the Spanish and English versions of the CDI for a bilingual child can cause their scores to fall below monolingual norms. Also issues with parents not understanding the instructions on the CDI led to

questionable results. One study included how some parents left confused comments on CDI forms that asked for clearer explanations of what they needed to do. Overall, most parents were able to administer the CDI in a confident manner. Face validity was high in all, but a few studies and parents felt for the most part, that their assessments accurately showed their child's language skills.

5.1.5 Disability

The CDI showed promise in being able to handle assessing the language skills of children with language or physical disabilities. Children with autism, down syndrome, clef-palate, and other language delays were able to have successful vocabulary reports completed. This is good news for those worrying about a child with a disability not being able to take the MacArthur-Bates CDI.

5.2 Limitations of the current body of research

This synthesis uncovered a few areas where the current research was lacking. One was the use of only one instrument to measure language outcomes. Another was the number of environmental variables that effected the results. Also, some sample sizes of the studies were rather small which weakens strength and validity. The small amount of Spanish and bilingual studies reported on compared to English also provides some limitations to the results of the study.

5.2.1 Instrument use and environmental limitations

Most of the studies only used the Macarthur-Bates CDI to measure the language outcomes of the participants. While the Macarthur-Bates is the focus of this synthesis, when trying to determine comprehension and production of language skills, it is helpful to incorporate multiple methods of language assessment. Another limitation arose in bilingual studies that had to do with

parents collecting the language data. Some parents may not know what words their child produces in both languages especially if the child uses a language at school which is different from home.

In all studies the CDI was administered at the home of the participants. While this is the actual setting where CDI's are normally completed, the environment cannot be controlled like a laboratory setting can. Distractions may have occurred in the home of the participants that were unknown by the researchers. The realistic setting was necessary for the studies to capture accurate representations of CDI use however, extraneous variables undoubtably effected the results.

5.3 Methodological issues

Many different styles were used by the researchers of these studies. When looking at many different studies there will undoubtably be differences in the way they are reported and designed. This resulted in hindering the strength of conclusions made after reviewing the results. Some participants were assessed multiple times over the course of a few months. Other were only tested once. Some studies did not disclose the ethnicity of all the participants while others provided detailed background information. Sample sizes varied across studies as well with a few reporting small numbers of participants which in turn can hurt the significance of results. There was variability on the tests and measurements used to interpret the data that was acquired from the participants.

Comparing results across the studies was hindered by the fact that not all studies included the same focus of measurement. Some only recorded total vocabulary, while others included both total and conceptual vocabulary. Some studies included information on mean length of utterance

or focused on production of closed class words. Some studies did not report on effect size which limits research significance as well.

5.4 Conclusions and ways to edit CDI

One way the MacArthur-Bates CDI can be altered is more descriptive sets of directions. There are no examples provided on the assessment for what it looks like when a child understands a word. Parents are left to make their own decisions which can lead to unreliable reporting. Corresponding video examples could be offered on the CDI website so that parents can follow along with the questions. Also, it may be beneficial for children from minority language families who are in L2 only school programs to have assistance from a caretaker or teacher when taking a version of the CDI in the L2 language. The child may speak very little of the L2 at home and parents may not know what words their child knows in the majority language. Having a caretaker or teacher from the majority language may help yield more accurate representations of the child's vocabulary.

This synthesis looked at how effective the MacArthur-Bates CDI was on Spanish monolingual, English monolingual, and Spanish and English bilingual speaking children from 2-36 months of age. The results show that using this assessment as a valid tool to assess vocabulary development is not without its limitations. The CDI can be influenced by various outside factors that may hinder the reliability of language outcomes. It is important to note that the reasons for exactly how variables, specifically SES and education effect the CDI are still unknown. If anything has been learned it is that future research on a larger scale should be done to provide more than just speculation on why these variables effect the CDI. We must remember studies have proven just because a child comes from a lower SES situation, doesn't mean their language skills are hindered. For now, parents using this assessment should understand that results may

not accurately display their child's language ability and any intervention decisions should still be corroborated by a professional physician.

VI. LIST OF REFERENCES

Armstrong, K., & Agazzi, H. (2010). Bayley-III Clinical Use and Interpretation. *Practical Resources for the Mental Health Professional*, 29-45.

Arriaga, R., Fenson, L., Cronan, T., & Pethick, S. (1998). Scores on the MacArthur Communicative Development Inventory of children from lowand middle-income families. *Applied Psycholinguistics*, 209-223.

Bates, E. M. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* , 85-123.

Brady, K., & Goodman, J. (2014). The Type, but Not the Amount, of Information Available Influences Toddlers' Fast Mapping and Retention of New Words. *American Journal of Speech Language Pathology* , 120-133.

Canestraro, B. (2014, March 11). *Expressive One-Word Picture Vocabulary Test (EOWPVT)*. Retrieved from Prezi: https://prezi.com/hmp-25lgrmix/expressive-one-word-picture-vocabulary-test-eowpvt/

Checa, E. G. (2016). The Composition of Early Vocabulary in Spanish Children With Down Syndrome and Their Peers With Typical Development. *American Speech-Language-Hearing Association*, 605-619.

Colona, K. (2019). Language Development Education.

Connery, A., Colbert, A., Lamb, M., Hernández, S., Martínez, M., Bauer, D., . . . Munoz, F. (2019). Receptive language skills among young children in rural Guatemala: The relationship between the Test de Vocabulario en Imagenes Peabody and a translated and adapted version of the Mullen Scales of Early Learning. *Child: care, health, and development* , 702-708.

Core, C. H. (2013). Total and Conceptual Vocabulary in Spanish-English Bilinguals From 22 to 30 Months: Implications for Assessment. *Journal of Speech, Language and Hearing Research*, p. 1637-1649. .

Dale, P. (1991). The Validity of a parent report measure on vocabulary and syntax at 24 months . *Journal of speech & Hearing Research* , 565-571.

de Diego-Lazaro, B. R.-I. (2019). Predictors of Vocabulary Outcomes in Children Who Are Deaf or Hard of Hearing From Spanish-Speaking Families. *American Speech-Language-Hearing Association*, 113-125.

Eadie, P., Ukoumunne, O. S., Prior, R., Bavin, E., Bretherton, L., & Reilly, S. (2010). Assessing early communication behaviours: structure and validity of the Communication and

Symbolic Behaviour Scales—Developmental Profile (CSBS-DP) in 12-month-old infants. *International Journal of Language and Communication Disorders* , 572-585.

Feldman, H., Dollaghan, C., T, C., Kurs--Lasky, M., Janosky, J., & J, P. (2000). Measurement propertiues of the MacArthur Communicative Development Inventories at Ages One and Two Years. *Child Development*, 310-322.

Fenson, L. D. (1994). Variability in early communicative development. 1-173.

Fletcher, K. P. (2005). Responsivnes and attention during picture-book reading in 18-24 month-old toddlers at risk. *Ealry Child Development and Care*, 63-83.

Foster-Cohen, S., J, E., Champion, P., & Woodward, L. (2007). Early Delayed Language Development in Very Preterm Infants: Evidence from the MacArthur-Bates CDI. *Cambridge University Press*, 655-675.

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018). A Cross-Language Study of Decontextualized Vocabulary Comprehension in Toddlerhood and Kindergarten Readiness. *Developmental Psychology*, 1317-1333.

Furey, J. (2011). Production and Maternal Report of 16- and 18-Month-Olds' Vocabulary in Low- and Middle-Income Families. *American Journal of Speech Language PAthology* , 38-46.

Gale, E. S. (2008). Symbol-Infused Joint Attention and Language Use in Mothers with Deaf and Hearing Toddlers. *American Annals of the Deaf*, 484-503.

Greenwood, C. W. (2010). The Early Communication Indicator for Infants and Toddlers Early Head Start Growth Norms From Two States. *Journal of Early Intervention*, 310-334.

Gulberson, M. R. (2011). Classification Accuracy of Brief Parent Report Measures of Language Development in Spanish-Speaking Toddlers. *American Speech-Language-Hearing Association*, 536-549.

H. Qi, C., & Marley, S. (2011). Validity Study of the Preschool Language Scale–4 With English-Speaking Hispanic and European American Children in Head Start Programs. *Topics in Early Childhood Special Education* , 89-98.

Heilmann, J. W. (2005). Utility of the MacArthur-Bates Communicative Development Inventory in Identifying Language Abilities of Late-Talking and Typically Developing Toddlers. *American Speech-Language-Hearing Association*, 40-51.

Holahan, A., & Costenbader, V. (2000). A comparison of developmental gains for preschool children with disabilites in inclusive and self-contained classrooms. *Topics in Early Childhood Special Education* , 224.

Houston-Price, C. M. (2007). Discrepancy between Parental Reports of Infants' Receptive Vocabulary and Infants' Behaviour in a Preferential Looking Task. *Journal of Child Language*, 701-724.

Hurtado, M. G. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish-English bilingual toddlers. *Bilingualism* , 189-202.

Jackson-Maldonado. D., M. V. (2013). Short-Form Versions of the Spanish MacArthur Bates Communicatice Developmental Inventories. *Applied Psycholinguistics* , 837-868.

Johnson, E. (2015). Debunking the language gap. *Journal for Multicultural Education*, p. 42-50.

Johnson, E. Z. (2017). Introducing the Language Gap. *International Multilingual Research Journal* , 1-4.

Johnson, L., Cook, M., & Kullman, A. (1992). An Examination of the Concurrent Validity of the Battelle Developmental Inventory c& Compared with the Vineland Adaptive Scales and the Bayley Scales of Infant Development. *Journal of Early Intervention*, 353-359.

Judith, K., & Maddux, T. (2014). Test Review. *Journal of Psychoeducational Assessment* , 577-580.

Levine, D. P.-P. (2020). Evaluating socioeconomic gaps un preschoolers' vocabulary, syntax, and language process skills with the Quick Interactive Language Screener (QUILS. *Early Childhood Research Quarterly*, 114-128.

Luyster, R. L. (2007). Characterizing communicative development in children referred for Autism Spectrum Disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Applied Psycholinguistics* , 623-654.

Mancilla-Martinez, J. P. (2011). Assessing the productive vocabulary of Spanish-English bilingual toddlers from low-income families. *Applied Pyscholinguistics* , 333-357.

Marchman, V. M.-S. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech Language and Hearing Research* , 983-997.

McDuffie, A. Y. (2005). Prelinguistic Predictors of Vocabulary in Young Children With Autism Spectrum Disorders. *Journal of Speech, Language & Hearing Research*, 1080-1097.

Mcleod, S. (2013). *Simply Psychology*. Retrieved from simplypsychology.org: http://www.simplypyschology.org

Mirenda, P., Smith, I., Vaillancourt, T., Georgiades, S., Duku, E., Szatmari, p., . . . L, Z. (2010). Validating the Repetitive Behavior Scale-Revised in Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 1521-1530.

Mullen, E. (1995). *Mullen Scores of Early Learning*. Retrieved from Psyche: v-Psyche.com

Pan, B., & Rowe, M. S.-L. (2004). Measuring productive vocabulary of toddlers in low-income families: concurrent and predictive validity of three sources of data. *Child Language*, 587-608.

Pearson, B. F. (1994). Patterns of Interaction in the Lexical Growth in Two Languages of Bilingual Infants and Toddlers . *Language Learning* , 617-653 .

Poulin-Dubois, D. B. (2013). Lexical access and vocabulary development in very young bilinguals. *The International Journal of Bilingualism*, 57-70.

Ramírez, N. L. (2019). Parent coaching increases conversational turns and advances infant language development . *PNAS*, 3484–3491 .

Roberts, J. B. (1999). Parents' Report of Vocabulary and Grammatical Development of African American Preschoolers: Child and Environmental Associations. *Child Development*, 92-106.

Roid, G., & Sampers, J. (2004). *Merrill-Palmer--Revised Scales of Development.* Wood Dale, IL: Stoelting Co.

Ryan, W. B. (2015). Measures of Attitudes towards Sexual Orientation . *Science Direct* , 719-751.

Scherer, N. (1999). The speech and language status of toddlers with clef lip and/or palate following early vocabulary intervention. *American Journal of Speech Language Pathology* , p.81.

Sperry, D. S. (2018). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Journal of Child Development* , p. 1-16.

Thal, D. O. (1999). Vaidity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech Language and Hearing Research* , 482-496.

Walker, D., Carta, J., Greenwood, C., & Buzhardt, J. (2008). The Use of Individual Growth and Developmental Indicators for Progress Monitoring and Intervention Decision Making in Early Education, Exceptionality. *A Special Education Journal* , 33-47.

Yoder, P. B. (1997). Stability of Maternal Reports of Lexical Comprehension in Very Young Children With Developmental Delays. *American Journal of Speech-Language Pathology* , 59-64.