

PREDICTING GLOBAL MEASURES OF DEVELOPMENT AT 18-MONTHS OF AGE  
FROM SPECIFIC MEASURES OF COGNITIVE ABILITY AT 10-MONTHS OF AGE

BY

Tasha D. Schmeidler

Submitted to the graduate degree program in Cognitive Psychology and the Graduate  
Faculty of the University of Kansas in partial fulfillment of the requirements for the  
degree of Master of Arts.

Committee:

John Colombo, PhD  
Chairperson

Committee members: Susan Carlson, PhD

Kathleen Gustafson, PhD

Date defended: 10/29/2009

The Thesis Committee for Tasha D. Schmeidler certifies that this is the approved  
Version of the Following theses:

PREDICTING GLOBAL MEASURES OF DEVELOPMENT AT 18-MONTHS OF AGE  
FROM SPECIFIC MEASURES OF COGNITIVE ABILITY AT 10-MONTHS OF AGE

Committee:

John Colombo, PhD  
Chairperson

Committee members: Susan Carlson, PhD

Kathleen Gustafson, PhD

Date approved: 10/29/2009

## Table of Contents

<b>I. Abstract</b>	4
<b>II. Introduction</b>	6
The Emergence of Executive Function	6
Neural Concomitants of Executive Function	8
<b>III. The Development of EF Components</b>	10
Problem-Solving in Infancy	10
Explicit Memory in Infants	11
The Bayley Scales as a Measure of Developmental Outcome	13
Predicting a Developmental Outcome: Specific versus global measures	14
Objectives	15
<b>III. Methods</b>	15
Participants	15
Measures	16
Willatts Problem-Solving Task Procedures ( <i>10 months</i> )	16
Elicited Imitation (EI) Task Procedures ( <i>10 months</i> )	17
Developmental Assessments	19
Bayley Scales of Infant Development Procedures ( <i>18 months</i> )	19
MacArthur-Bates Communicative Development Inventory ( <i>18 months</i> )	19
Statistical Plan	20
<b>IV. Results</b>	21
Factor Analysis	22
Regression Analysis with Standardized BSID-II Scores	23
Regression Analysis with Factor Scores	23
<b>V. Discussion</b>	24
Factor Analysis Findings	25
Predictive Properties of the Willatts Problem Solving Task	26
Predictive Properties of the Elicited Imitation Task	27
MBCDI-WS	27
Can Specific Measures Predict Global Measures?	28
<b>VII. References</b>	31
<b>VI. Appendices</b>	35
A. Table 1: Scoring Procedure for Willatts Problem Solving Task	36
B. Table 2: Descriptive Statistics for 10-Month and 18-Months Subjects	37
D. Table 3: Factor Correlations	38
E. Table 4: Factors extracted from BSID-II MDI subjects' item responses	39
F. Table 5: Loadings of Individual BSID-II Test Items onto Each Factor	40
G. Table 6: Formula for Estimating Factor Scores for Each Subject	41
H. Table 7: Using 10-month Tasks to Predict 18-month BSID-II Scores	42
I. Table 8: Using 10-month Tasks to Predict 18-month Factor Scores	43
J. Table 9: Correlations among each of the predictor variables	44

### *Abstract*

Composite scores from standardized tests taken in infancy have been shown to offer modest prediction to cognitive skills in later childhood. One possible reason for this is that early manifestations of mental development in infancy are not the basis for cognition later in childhood. One alternative hypothesis, however, is that the aggregation of composite scores on infant standardized tests obscures the measurement of specific skills that would be more predictive of meaningful outcomes.

In this study, we sought to determine if (a) it was possible to extract measurement of specific skills from a standardized infant assessment, and (b) whether the indicators of those skills might relate to performance on specific tasks administered in the laboratory. To accomplish this, we derived factors from an item-level analysis of an 18-month sample of subjects' MDI responses on the Bayley Scales of Infant Development, 2<sup>nd</sup> Edition (BSID-II) and tested to determine whether these specific factors predict to, or are predicted by, specific cognitive measures given to the same sample subjects at 10-months of age. In addition, we sought to determine whether the use of factor scores as dependent variables would demonstrate better predictability than the MDI Standard Score of the BSID-II as a dependent variable.

Three significant factors were derived from the 18-month-old subjects' responses to the MDI of the BSID-II: Based on item content, we named these Expressive Language, Receptive Language, and Object Manipulation. Factor scores were then calculated for each subject and entered into a multiple regression model to examine how the scores of the specific measures of cognitive abilities improved predictive ability to factors derived from the standardized development test. Scores from specific measures of problem-solving and explicit memory given to 10-month-old infants were utilized to examine possible predictability to a standardized developmental test, the BSID-II, given to the same sample ( $n=109$ ) of infants at 18-months of

age. Predictability of a specific measure of development to a global measure of development was promising in some instances, and the use of factor scores did improve the predictability for some outcomes. Implications of these particular analyses, including specific differences between these subscales and other's subscales that have been attempted previously are discussed. We also compared the factors to the word count and sentence length portions of the McArthur-Bates Communicative Development Inventory (MBCDI), which yielded promising correlations.

## *Introduction*

### *The Emergence of Executive Function*

From the viewpoint of cognitive psychologists, executive function (EF) consists of a number of cognitive strategies, which are involved to reach an end goal. Neisser (1967), referred to EF as “executive control,” which leads to several cognitive processes working together during goal-oriented problem solving. Posner (1978) suggested the presence of a “central processing system” that controls inhibition of behaviors that are not necessary for a specific task, but which leads to the ability to create effective strategies to reach problem-solving goals. As well, Developmental psychologists have added that planning and self-monitoring behaviors are involved in improved performance in cognitive contexts (Welsh, 1988).

There is much agreement among researchers today that EF is comprised of several cognitive components that include the ability to inhibit irrelevant information and allocate attention, as well as working memory and cognitive flexibility. Diamond and Goldman-Rakic (1985) used data that had been gathered in animal models and attempted to integrate this information with known outcomes from behavioral testing in developmental psychology. They used a classic task associated with EF, the delayed-response task. In the delayed-response task, the infant is seated in front of two empty wells. An object of interest, such as a small toy, is hidden in one of the wells and the infant is allowed to search for it after a brief delay of up to 10 seconds. Once the infant is successful in finding the object of interest, the toy is hidden in the other well on the opposite side. In the task, experimenters have found on several occasions that infants under 12-months of age make the classic “A-not-B error” that was observed by Piaget in the 1950s; that is, the infant will reach back to the initial side in their search for the object of interest (location “A”), rather than the side on which the object has been hidden (location “B”). Piaget (1954) stated that this error occurred because infants had poorly established concept of objects, poor motor-based responses, and a failure to update their referenced spatial coding scheme. Diamond and Goldman-Rakic (1985b) found that infant

monkeys and infant humans made this same mistake on a delayed-response task. This led to a model of executive functioning that included relating temporally or spatially separated information and inhibiting prepotent responses (i.e., responses that have been previously-established, either through natural tendencies or prior learning).

Despite the presence of a considerable literature on tasks that are presumably EF-related during infancy, there is a noticeable gap in the EF research for children between the ages of one- and three-years; this is especially true between the ages of one and two years. McGuigan and Nunez (2006) related this to the complexity of EF task tasks that have been developed thus far: the tests used in infancy are too simple for toddlers, but that other tasks are too complicated and have too many steps for a toddler to complete successfully. In an attempt to demonstrate the development of EF within this age group, McGuigan and Nunez (2006) used a modified version of the detour reaching task. In the infant version of this task, a desired object is placed in a transparent box within the infants reach; however, the closed side is placed directly in front of the infant, requiring the infant to reach around to another side to retrieve the desired object. Young infants typically reach directly in front of them to retrieve the object, but at 11-12 months of age, the infant is usually able to inhibit the prepotent response of reaching directly in front of them to retrieve the object and reach around the transparent barrier to retrieve the desired object (Diamond, 1990). Preschool versions of the task have also been developed. These versions often require the child to perform another action (such as flip a switch to move another barrier), before he/she can successfully reach around the transparent barrier to retrieve the desired object. This version of the detour reaching task has been successful measures of EF in preschool-aged children because it not only requires the child to inhibit the prepotent response of reaching directly for the object of interest, but it also requires the child to keep in mind the series of steps needed before the child can retrieve the object. McGuigan and Nunez (2006) found that toddlers could perform successfully on the detour reaching task if the box was opaque rather than transparent; thus, the extra step was causal, not arbitrary. The opaque box

presumably lowered inhibitory control demands, since the object could not be seen, and the causal step was more easily understood by toddlers than a step that didn't seem to make sense with the task.

Executive Function improves gradually for typically developing young children. EF would be characterized as poor for children with developmental disorders such as Attention-Deficit/ Hyperactivity Disorder, or Autism (Powell & Voeller 2004). Children with poor executive function demonstrate difficulties with sustained attention tasks, dealing with novelty, transitioning from one activity to another, distractibility, controlling impulsive responses, and regulating emotional states (Powell & Voeller 2004). Diamond (1988) relates a child's poor performance on tasks requiring executive function to poor inhibitory control. She hypothesizes that EF represents an occasion where cognitive control is required but is not sufficient to reach one's goals (Diamond 2006). A key aspect to Diamond's model of EF is the involvement of inhibition of an automatic tendency to act on information held in the mind, or as she refers to it, "the conjunction of working memory and inhibition."

More recently, Colombo and Cheatham (2007) proposed a model of endogenous attention that posits the basis of EF in early childhood as the integration of attention with working memory. In this model, the emergence of endogenous attention is related to the concurrent development of specific memory functions toward the end of the first year of life. They explain that the development of the ability to voluntarily hold attention is due to the ability to link working memory with the fundamental attentional functions. These models all possess the same common theme: incorporating working memory with the voluntary deployment of attention, intention, and planning in the emergence of integrated higher order cognitive functions.

#### *Neural Concomitants of Executive Function*

Dramatic changes occur in the brain during infancy and early toddlerhood, including increased gray and white matter volume, rapid synapse formation and subsequent pruning of synapses in the prefrontal cortex (Casey et al., 2000). Synaptic formation in the prefrontal

cortex occurs throughout infancy and toddlerhood, reaching a peak level of synaptogenesis at about 24-months of age (Huttenlocher, 1994). Past research involving brain imaging in typically developing humans and those with deficits has strongly associated the dorsolateral and ventrolateral prefrontal cortex in the mediation of many forms of EF (Diamond 2006).

Diamond and Goldman-Rakic's (1985b) studies of a delayed-response task in infant humans and infant monkeys suggested that EF was localized to the frontal lobes. Adult and infant monkeys with frontal lobe lesions consistently made the typical A-not-B error response in the delayed-response task, while this deficit was not demonstrated when adult monkeys experienced lesions in other areas of the brain.

Various aspects of EF have been related specifically to areas of the prefrontal cortex on numerous occasions throughout the last few decades. In one of the first known attempts to integrate the prefrontal cortex with EF, Fuster (1985) integrated the known functions associated with the prefrontal cortex to EF in three ways: 1) the temporally retrospective function of working memory, 2) the temporally prospective function of anticipation of events, and 3) interference control and suppression of behaviors not needed to reach the end goal.

Baddeley and Hitch (1974) proposed that the dorsolateral prefrontal cortex (DLPFC) was involved in the "central executive", which controls and manipulates information being held in mind, and also regulates and integrates cognitive abilities. The DLPFC receives information from circuitry connected to the temporal and parietal lobes of the brain, which might integrate information about objects and their meanings, location of objects, and the emotional status of others (Powell & Voeller 2004). The DLPFC has been specifically linked to EF because it specifically mediates attention and focus (i.e., affects distractibility), and allows for flexible shifting of specific cognitive abilities to others (Powell & Voeller 2004): abilities that are central to effective control of higher order cognitive functions. Researchers today are in agreement that EF is very diverse in nature and involves multiple cognitive processes, but with its fundamental basis in the prefrontal cortex (see Weibe 2004).

### *The Development of EF Components*

*Problem-Solving in Infancy.* Problem-solving involves goal-oriented behaviors such as selecting actions that are appropriate to the goal, correcting errors in one's efforts to achieve the goal, and stopping when the goal is reached (Bruner, 1981). In infants and toddlers, the nature of how these steps are taken to reach an end goal is of interest.

The ability to hold an object of interest in mind for a longer period of time improves at approximately 8 to 9 months of age (Diamond 2006). At this age, it has been noted that an infant begins to act on one object to get another in order to successfully retrieve objects of interest (Piaget, 1954). Holding an object in mind and acting on an object to retrieve another object are key aspects to problem-solving behavior.

Problem-solving behaviors in infancy are usually best demonstrated by means-end tasks. Tasks are considered "means-end" if they involve a sequence of ordered steps that are required to reach or attain a goal. Sometimes these steps include a type of obstacle that must be overcome before the goal can be reached. According to Willatts (1999), means-end behavior in infants includes the following characteristics:

- (a) the infant has a definite goal and is persistent in achieving it; the infant completes the initial step that achieves the subgoal of setting up the preconditions that permit the final goal to be achieved by direct action; and,
- (b) the infant produces behavior that is appropriate for accomplishing the first step, resulting in an overall sequence that is organized and goal-directed.

One particular task that has been used to assess problem solving in infancy combines each of these components to take a comprehensive look at problem-solving. Developed by Willatts (1984b), this "means-end" task involves a trial where the infant must retrieve a toy hidden under an opaque cloth, a trial where an infant must pull a support to retrieve a distant object, and finally a trial where the infant must remove a barrier to retrieve a toy. These three components from the baseline trials are then combined, and the infant attempts follow the sequence of three

steps to retrieve the toy. Scores are based on retrieval of the toy and the intentionality of that retrieval. Improvements in intentionally completing means-end tasks improve from 6 to 8-months of age (Willatts, 1999). Improvement in means-end tasks also improves in later infancy as well (Willatts 1984b). Willatts noted that infants at 6 months of age did not intentionally search for the toy. Once infants were 8-months-old, however, they appeared to search for the object more intentionally, suggesting the ability to have goal-directed behaviors. Willatts hypothesizes that infants may possess the knowledge that they have to complete a certain action to reach their goal of retrieving a toy, but that they may not necessarily know the precise action to take to reach this goal (1999).

Problem-solving strategies in toddlers have been addressed by Chen and Siegler (2000). They noted that when toddlers are given tools to solve a problem, they are usually not effective in reaching their goals if left to figure out a problem on their own. However, if they are given a demonstration of how to use a tool to solve a problem or a verbal hint, they can be successful. Chen and Siegler (2000) also noted that toddlers up to 3 years of age have difficulties in selecting tools on their own for strategic problem-solving.

Effective problem-solving in infants and toddlers depends on several aspects of EF, including the ability

- a) to plan
- b) to hold information in mind for extended periods of time to work through a problem
- c) to inhibit prepotent responses that may interfere with successfully solving a problem efficiently.

*Explicit Memory in Infants.* Explicit (declarative) memory involves rapid recall of items from past experiences, such as important dates, names, or events. Piaget (1952) believed that infants and toddlers lacked the capability to form symbols or representations of previous experiences. Measuring recall of events in preverbal infants required researchers to create non-verbal tasks that could demonstrate explicit memory in infants. Meltzoff (1985) used such

tasks that demonstrated toddlers at 24 months of age could successfully remember from a task modeled for them 24 hours earlier.

In a deferred imitation task, a child is presented with an object or objects that they become familiar with. The experimenter then demonstrates a goal or action that can be attained by arranging those items in a particular sequence. After a delay, the child is then asked to demonstrate the action they were shown previously. Additionally, Meltzoff (1988) demonstrated that infants as young as 9-months of age were able to defer imitation of a single action for 24-hours, and 50 percent of the infants at this age were able to remember two or three of the actions involved in a sequence. Their observations were consistent with the implications of memory from the conjugate reinforcement paradigm that originated in the 1960s (Rovee-Collier 1997).

Bauer and Shore (1987) developed a paradigm that they called *elicited imitation* (EI). Elicited Imitation is analogous to deferred imitation in that it requires the infant to remember a sequence of events involving objects but here, there is no delay introduced, and the infant is asked to immediately demonstrate the actions modeled by the experimenter. Elicited Imitation and Deferred Imitation are thought to be effective ways to measure the abilities of an infant to have a mental representation that leads to an explicit memory (Bauer 2004).

It is currently held that long-term recall of events emerges before an infant can express recall of past experiences verbally, are abilities that emerge in the first year of life and become more reliable over the second year of life (Bauer 2002). EI tasks can be used to demonstrate abilities of EF, including

- (a) the inhibition of prepotent responses involved in remembering how to complete a sequence of events,
- (b) planning how to complete a sequence of events, and
- (c) higher order memory functions, such as working memory.

Bauer (2002) adds that retrieving memories after a delay (i.e., long-term recall) require intact connection between the temporal and prefrontal structures of the brain; prefrontal structures of the brain that are also associated with executive function.

### *The Bayley Scales as a Measure of Developmental Outcome*

One of the most widely used measures of developmental function is the Bayley Scales of Infant Development (BSID). The BSID was developed by Nancy Bayley in 1969, and originated from three other tests that came about in the 1930s: The California First Year Mental Scale (Bayley, 1933), The California Preschool Mental Scale (Jaffa, 1934), and the California Infant Scale of Motor Development (Bayley, 1936). The BSID includes items that purport to measure mental and motor abilities in infants and toddlers from one- to 42-months of age. The test also includes a behavioral rating scale that is filled out by the examiner after the assessment, which encompasses orientation to the examiner, emotional regulation, and behavioral ratings. A second version (BSID-II) was released in 1993, and a third edition of the BSID was released in 2005; the current study involved the BSID-II.

The BSID-II is used to document the progress of development at specific ages and in so doing, it includes the assessment of some EF abilities, including: planning how to execute a correct response on a test item, inhibiting impulsive responses, and regulating emotional during difficult or frustrating test items. Children who are diagnosed with disorders associated with dysfunction of EF (e.g., ADHD, schizophrenia, or PKU), are more likely to perform poorly on a standardized test of cognitive ability or intelligence than a typically developing child (Blair 2006); thus, EF might be expected to be reflected in global measures of development or cognitive function.

One criticism of the BSID is that it is a global measurement that might be less sensitive to variability in individual behavior (see Bandura 1981). It has long been noted that the BSID is a poor predictor of later developmental outcome, particularly during the first year. McCall (1981b) reviewed the literature on the BSID and other infant measures and concluded that the predictive

validity of these measures was relatively poor, except in extreme cases of lowered cognitive abilities.

There have been some successful attempts to use the BSID as an outcome from cognitive tasks (Rose, Feldman, & Wallace, 1992); there have also been several unsuccessful attempts (Messer, McCarthy, & McQuiston, 1986; Hack et al., 2005) to demonstrate association between specific measures of cognition and later performance on the BSID. Along with these attempts to demonstrate predictability from or to standardized tests, the idea of deriving subscales from such tests to isolate specific cognitive abilities came about. It was suggested that subscales demonstrating specific cognitive abilities from particular clusters of items on standardized tests might be more effective than simply the overall scores from these tests.

There have been at least two attempts to derive subscales from the BSID. Dale et al. (1989) attempted to derive subscales from the mental scale of the BSID, but these were primarily concerned with specific aspects of language (Molfese & Acheson, 1997). The Kohen-Raz system (1967), which proved quite useful for some time and was relatively widely used to better predict later developmental outcomes (see Molfese & Acheson, 1997), identifies five different subscales from the first version of the BSID: Eye-hand ability, manipulation, Object Relations, Imitation-Comprehension (spoken language abilities), and vocalization-social (verbal expressiveness). We are unaware of any attempt to derive subscales with the BSID-II. Another important note is that the Kohen-Raz system has been used primarily to predict later outcomes from one global test to another global test (Molfese & Acheson, 1997), rather than to predict to a specific measure of cognitive ability to another global test. Continuity of cognitive development might be demonstrated if specific measures of cognitive development would predict to a later developmental outcome.

#### *Predicting a Developmental Outcome: Specific versus Global Measures*

The Mental Development Index (MDI) of the BSID-II is a composite assessment that includes areas such as social skills, concepts of numbers, and development of habits (Lichtenberger,

2005). The BSID-II implicitly measures specific abilities involving EF, including memory-learning, problem-solving, verbal abilities, abstract thinking these abilities are required to complete the items of the BSID-II. The Willatts Problem-Solving Task and Elicited Imitation both give a measure of abilities that could be precursors to later EF. It would be of interest to examine if developmental outcomes of these specific measures of cognitive abilities better predict later aptitude on the BSID-II if subscales of the BSID-II mental scale could be successfully derived.

### *Objectives*

The purpose of this analysis is to examine the ability of the Willatts Problem-Solving Task and the Elicited Imitation in a cohort of children, both of which are given to this particular sample at age 10-months, to predict a developmental outcome on the BSID-II at 18-months of age. To accomplish this, not only will we look at the raw and standardized scores of individual subjects on the BSID-II in comparison to the performance on the tasks with specific measures, there will be an attempt to derive subscales within the BSID-II mental scale as Kohen-Raz (1967) did for the BSID, and to examine whether the predictive ability of the specific measures improves with the calculations from subscales. We will also make comparisons between the derived subscales and aspects of the child's language development via the MacArthur-Bates Communicative Development Inventory (MBCDI). Finally, comparisons will be made to see if the current subscales resemble those created for the original BSID by Kohen-Raz, keeping in mind that the Kohen-Raz subscales were created to encompass test information from several age groups rather than just one specific age group.

### *Method*

#### *Participants*

The study used a convenience sample of 123 infants was recruited for a randomized, double-blind, clinical trial investigating the effects of infant formula supplemented with long-chain polyunsaturated fatty acids on visual and cognitive development. For this particular study,

women were recruited and consented in a hospital visit after they had delivered a term newborn to participate in the study until the subject was 18 months of age. Women were approached only if they had decided to not breast-feed and had already given their infants formula supplemented with different levels of docosahexaenoic acid (DHA), a long-chain polyunsaturated fatty acid, or regular formula without supplementation (control) from birth to 12-months (or until the infant was weaned from formula). During this time, infants were seen by a registered dietitian, an electrophysiologist for a visual acuity assessment, and cognitive scientists at ages of 6 weeks, 4, 6, 9, 10, 12, and 18 months of age. At each visit a 24 hour diet recall and anthropometrics were obtained as well as measures of visual and cognitive development. Subjects were compensated \$50 per visit for their time. For these particular analyses, however, the DHA levels data is not utilized. Because this was an ongoing clinical trial, experimenters are still blind to this information; only normative data from the cognitive testing was included.

At 10 months of age, infants were administered the Willatts Problem-Solving Task followed by an Elicited Imitation Task, both which have been described briefly above. At 18 months of age, infants were administered the BSID-II. To have a better understanding of the types of analyses that will take place, descriptions of the tasks and how they are subsequently scored are given below.

### *Measures*

*Willatts Problem Solving Task Procedure (10 months).* Each 10-month session was videotaped in its entirety for later coding purposes. A video camera was placed behind the experimenter at an angle which allowed a clear view of the infant being tested and the area of the table right in front of them. During the 10-month session, infants were first seated on their mother's lap at a medium-sized table across from the experimenter. Parents were instructed to let the infant figure out the task on his/her own with no help from them. The infant was first familiarized with the objects associated with the task - a small black cloth, a larger blue and white checkered cloth, a soft blue barrier, and an object of interest (which was a small clear

rattle with multi-colored beads inside of it) for 20 seconds each. The infant was then given 30 seconds on each of the following problem-solving trials as a baseline: The rattle was hidden under the black cloth for the infant to find, then placed at the end of the checkered cloth that was furthest away from the infant for them to pull to themselves, and finally, the soft barrier was placed in front of the toy requiring the infant to move it out of the way to retrieve the toy. If the infant did not retrieve the toy on the first trial, they were given a second trial. Scores given were fail (0 points), object retrieved without clear intention (1 point), object retrieved with clear intention (2 points). The infant's level of intention was determined by trained coders, and was determined reliable if two coders were in agreement on the infant's level of intention. If the infant did not retrieve the toy on one of the three trials, the task was terminated. If the infant passed on each of the three baseline trials, the infant was then given the challenge to complete a series of the three baseline trials with the final goal of retrieving the toy: That is, the infant had to first remove the barrier, and then pull the checkered cloth to bring the items closer to them, and finally remove the black cloth in order to find and retrieve the toy. Infants were given 30 seconds to complete the series of events, with the final goal of retrieving the object of interest. Scoring of these sessions was done after the session from a video recording of the session. The task is scored based on nine criteria that take place during each step that eventually leads to the retrieval of the object of interest (see *Table 1*). The points are totaled for each attempt, and an average score is taken from all three attempts to retrieve the object of interest.

*Elicited Imitation (EI) Task Procedure (10 months)*. The EI task immediately followed the Willatts Problem-Solving Task during the 10-month testing session. The infant was given a warm-up toy with a dump-truck and two colored-blocks, and instructed to put the blocks in the truck, and then dump them out. After the warm-up, the infant is presented with one of four testing events (labeled Event A, B, C, or D and presented in that order) to first explore on their own for approximately two minutes. During this time, the infant receives no assistance or interference from the experimenter, and is simply allowed to explore the object(s) on his/her

own. Each of the test events involves an item or a group of items that when put together in a particular sequence creates one unitary item with a specific purpose. After the infant is allowed to explore the presented items, the experimenter then models the sequence of actions required to complete the event. The infant is then either given the items back immediately to attempt, or given the items back after a brief delay to attempt. Each session involves Event A as immediate recall, Event B as delayed recall, Event C as delayed recall, and Event D as immediate recall. At the end of event B, a stopwatch is started to create a minimum 12-minute delay before the *delay events* (B and C) are presented again. Event A is then presented again at the very end to see if there is any recall for that particular event. During the time remaining in the delay after the infant has seen each of the events, the infant is allowed to play with different toys not involved in previous EI task.

The four possible events were counterbalanced among the subjects and are as follows (More details of these events can be found in Bauer (2004)):

**Find the Bear:** The participant is presented with a wooden box with an attached door. To “open the door” and reveal a toy bear, the participant must first “pull the lever.”

**Pop the Puppy:** A large box that requires the participant to “open the gate” and then “push in a block” to cause a stuffed puppy to “pop” out of a box.

**Make a Turtle Slide:** The participant is required to “build a ramp” by folding two wooden pieces at about a 45 degree angle that are joined with a small hinge, and then “slide” a plastic turtle down the slide that was created.

**Make a gong:** The participant is required to “hang up” a metal plate on a plastic base and “ring it” by striking the metal plate with a small, plastic hammer.

The session is coded from a recording of the infant performing the task. If the infant shows clear intent to complete individual steps to reach the final goal, they are given a point for completing each item. The infant also receives a separate score for completing items in the exact order that they were modeled by the experimenter. For example, if the infant is asked to

recall the event “Make a Gong”, and he/ she completes the event exactly as modeled, the infant will receive two points for completing both components to reach the final goal state, and then one point for completing them in the proper order. If the infant simply remembers the step “Ring It!”, he/she receives one point for recalling that particular component, but none for the omitted component, nor are any points received for proper order completion of the item. The most points that an infant can achieve in this task is two points for the components, and one point for the pairs.

### *Developmental Assessments*

*BSID II (18-Months)*. During the 18-month visit, infants were given the BSID-II. The BSID-II was given in a large testing room, which allowed the child to have plenty of room to perform effectively on the mental scale items as well as the motor scale items. The experimenter was seated at a medium-sized table with the child in a child’s booster seat directly across from them. A parent was seated next to the child, and instructed not to help the child with any of the testing items, or give hints on how to perform on each item. The Mental Scale (MDI) was given first, followed by the Motor Scale (PDI), and then the Behavioral Ratings Scale (BRS) was filled out by the experimenter right after the child finished testing. Experimenters followed a specific order of administering items on the mental and behavioral scales, and scored each item as the test progressed. Items were administered by each experimenter according to the manual that came with the test (see Bayley 1993).

*MacArthur-Bates Communicative Development Inventory (18-months)*. At the 18-month visit, the subject’s primary caregiver was also given the *MacArthur-Bates Communicative Development Inventories: Words and Sentences*, or *CDI-WS*. (Fenson, et al., 1993). This is a parent-report item for toddlers, and is designed to assess vocabulary knowledge and grammatical skills. Parent reports have been used for decades (Fenson et al. 1993), however, they are sometimes criticized as a primary form of assessment because a parent may not provide an accurate report. Fenson et al. (1993) noted that parents may not have specialized

training in language development and may not be sensitive to subtle aspects of language structure and use; On the other hand, a child's behavior in day-to-day activities when they are not visiting the lab is more likely to represent of his/her linguistic abilities.

The MacArthur-Bates CDI-WS was given to the parent just after the BSID-II was administered. Parents are instructed to fill in the circle next to words that their child is currently spontaneously producing; Researchers stressed to parents during the visit to not include words that the child is simply repeating, or to not include words that the child may know, but is not currently saying. The inventory captures if the child is combining words and the length of their longest three sentences. These three are then averaged to find each child's "average sentence length" as reported by the parent during their 18-month visit to the lab. This can take the parent approximately 20-40 minutes depending on the child's current level of development (Fenson et al. 1993), and a researcher is available in the room if the parent has any questions or needs any clarification as they go through the inventory.

### *Statistical Plan*

We had two purposes for the analyses of these data. The first was to determine whether it was feasible to create reliable subscales of the BSID-II. To create factors of the BSID-II, each subject's response to each item on the mental scale was entered into a database. The scores were entered initially as "1" for a correct response, "0" for an incorrect response, and "R" for a subjects refusal to complete an item, and "." for missing data due to a subject either missing the 18-month visit. The data were then entered into a factor analysis that was appropriate to the dichotomous (yes/no) data set to determine if any of the items fell into particular categories, or cohered, with one another.

Second, we sought to determine whether the 10-month measures of specific cognitive abilities could predict factors extracted through the procedures described above from the BSID-II. Using subjects' scores on the Willatts Problem-Solving Task and the Elicited Imitation Task, regression analysis was used to determine if these tasks were able to predict how well subjects

performed on the BSID-II, using standardized scores and factor scores derived from the analysis described in the preceding paragraph.

### *Results*

The original sample size recruited for this study was 123 participants. Three participants were excluded from the final analysis due to possible developmental delays ( $n=1$  unconfirmed auditory problem;  $n=1$  unconfirmed neurological problem;  $n=1$  confirmed delayed motor development). If a participant failed to complete the 10 month visit but did complete the 18 month study visit, or vice versa ( $n=27$ ), their scores and item responses were estimated using the SAS (v 9.2) statistical procedure for multiple imputation, PROC MI. These scores were then entered into the database before the factor analysis was completed. This was done in order to make better use of the data we had from all participants, and make the factor analysis more accurately reflect our particular sample. If the participants did not complete either the 10-month session or the 18-month session, they were excluded from the analysis ( $n=15$ ).

The final analysis included a sample size of  $n=104$ ; one additional subject was dropped from the final analysis due to an extremely low MDI score compared to the rest of the sample (MDI = 73). One of the experimenters noted in the database that the subject did not speak at all during the session, and that this may have not accurately represented the subject's typical behavior. Of the final sample, 33 of these subjects had imputed 10-month data, and 30 of these subjects had imputed 18-month data. *Table 2* contains a summary of the samples scores on the tasks. It should be noted that on the Elicited Imitation Task, only whole number scores of 0, 1, or 2 can actually be attained; however, the Willatts Problem-Solving Task, a score of a fractional number is plausible.

Based on administration of the BSID-II at the 18-month session, the subjects' responses were entered into a database as "1" for a response with credit received on a test item, and "0" if no credit was received on a test item. Because the "correct" and "incorrect" responses were entered as "1" or "0" respectively, it was necessary to create a matrix of tetrachoric correlations

to accurately reflect the responses on the BSID-II in the factor analysis. Tetrachoric correlations are entered into the factor analysis in these instances because it is the recommended method of factor analyzing dichotomous data sets (Kubinger 2003; Parry & McArdle 1991).

### *Factor Analysis*

Once the matrix of tetrachoric correlations was derived from the raw BSID-II item-level data (see *Table 3*), CEFA Tool 2.00 (2004), a comprehensive exploratory factor analysis software, was utilized to perform a factor analysis. The ordinary least squares, or OLS, method was used in this instance because this particular matrix of tetrachoric correlations was “not positive definite”, meaning the matrix lacked normality and contained negative EigenValues. This occurs often when a tetrachoric correlational matrix is derived because in most instances, this type of correlational matrix leads to negative EigenValues, which leads statistical software to interpret this as a matrix that violates normality requirements. OLS works well with tetrachoric correlational matrices because OLS does not require normality in the matrix as other types of factor analytic methods do. To increase interpretability of the factor loadings, an orthogonal-varimax rotation was then applied. It has been noted in the literature that orthogonal rotation may not be the best method for use in psychological research because, by definition, the method it does not allow the factors to be correlated. Many argue that this would not be a viable assumption with psychological data. However, there are two reasons for its use here. First, theoretically, we were looking to extract factors from the Bayley composite based on the underlying notion that the composite was obscuring potentially important effects by combining data from distinct and independent behavioral domains; thus we were, by definition, looking for orthogonal factors. The second is a practical consideration; Guilford (1964) suggests that if the factor correlations are essentially zero, then the use of the orthogonal rotation is not inappropriate. In this particular case, correlations between factors were very close to zero (see *Table 3*).

From the final matrix, three factors were successfully derived from the subjects' responses on the BSID-II items 97-127, which are administered to 17 to 19-month-olds (see *Table 4*) using the methods described in the previous paragraph. A "flat loading rule" was applied to sort out which item loaded onto what factor. Most items clearly loaded onto a particular factor, while the loadings for the remaining factors would be either close to zero or negative (see *Table 6*). The first factor clearly reflected expressive language with Items 106, 109, 113, 114, 117, 121, and 127; the second factor involved items 99, 101, 108, 109, 120, 111, 118, 122, 124, 125, and 126, which required receptive language abilities; and, the third factor included test items 105, 112, 115, 120, 123, which involve skill in object manipulation. Descriptions of these items can be found in *Table 5*, along with their corresponding factor loading. These three factors will also be discussed more in depth in the Discussion portion of this document.

From the factor loadings that resulted from the factor analysis, factor scores for were calculated for each subject based on their item responses on the BSID-II for each of the factors 1, 2, and 3 using the regression method for calculating factor scores (see *Table 6*). These factor scores were then used as final predicted values (dependent variables) from the Willatts Problem-Solving Task and the Elicited Imitation Task given at 10-months of age in the regression analysis.

#### *Regression Analysis with Standardized BSID-II Scores*

A regression analysis was used to determine if the scores on the Willatts Problem-Solving Task and the Elicited Imitation Task given at 10-months of age were effective predictors to the MDI Standard Scores of the 18-month-old subjects on the BSID-II. *Table 7* provides subjects' score information with corresponding descriptive statistics.

*Regression Analysis with Factor Scores.* It was also of interest to see if the predictive value of the 10-month tasks increased when they were correlated with factor scores derived from the BSID. This analysis addressed the question of whether there was an underlying proficiency that

could be detected at early age that would help researchers predict how well a child would perform on later standardized tests.

In order to see if the tasks given to subjects at the 10-month visit could predict outcomes of our sample at 18-months of age, a multiple regression technique was applied to the data with the independent variables being the subject's Willatts Problem-Solving Task Score and Elicited Imitation Task Scores for immediate and delayed recall, and the dependent variables being each of the three factors derived for each subject from the factor analysis.

The analysis suggested that there was only slight evidence for the predictability of the tasks given at 10-months of age to the BSID-II factor scores derived at 18-months. As can be noted from *Table 8*,  $p$  never reaches a significant level in each of these individual regression analyses, except in the case of the Willatts Problem Solving Task as a predictor to Object Manipulation ( $p=0.05$ ). Implications of this outcome will be discussed in a later section of this document. Interesting correlations were also found between the derived factors and language development data from the MBCDI: Expressive Language correlated significantly with sentence length, and Receptive Language correlated significantly with the number of words a child was saying at the time of their 18-month visit. These correlations were stronger and more significant than the relationship between the components of the MBCDI and the MDI composite score alone. These findings will be discussed in the section below.

### *Discussion*

The first goal of this analysis was to determine whether subjects' item responses on the BSID-II at 18-months of age could be successfully parsed into coherent factors. This goal was achieved by first creating a matrix of tetrachoric correlations from the subjects' item responses on the 17-19 month section of the MDI portion of the BSID-II. Three factors emerged from a factor analysis using the OLS method, and an Orthogonal-Varimax rotation was applied to these factor loadings to increase the interpretability. Based on their content items, the three factors were labeled Expressive Language (items that require the subject to produce speech or

language), Receptive Language (items that require the subject to understand words to attempt tasks), and Object Manipulation (items of the BSID-II that involve placing pieces correctly on puzzle boards, building towers of blocks, and successfully retrieving a toy from under a clear box with an opening on the left or right side).

### *Factor Analysis Findings*

One of the first questions proposed about the derived subscales is how they differ than those derived for the first edition of the Bayley Scales of Infant Development by Kohen-Raz (1967). As mentioned earlier, Kohen-Raz found five different factors: Eye-hand, Manipulation, Object Relations, Imitation-Comprehension (spoken language abilities), and Vocalization-Social (verbal expressiveness); in this particular analysis, three significant factors were extracted: Expressive Language, Receptive Language, and Object Manipulation.

Similar to Kohen-Raz, this analysis revealed a strong language component, as two of the three factors involving this skill. The other factors retained by the Kohen-Raz scalogram analysis are similar to the “Object Manipulation” factor found in the current analyses; however the current analysis grouped the items together as a whole. This could be due to differences in technique used to derive the subscales (scalogram vs. factor analysis). The Kohen-Raz method also did not utilize a factor analysis based on tetrachoric correlations, so it is possible that the current method of analysis may have picked up on underlying abilities not recognized previously.

These underlying abilities may be specific to 18-month-olds as well, and differences in our sample should be noted. Kohen-Raz found factors based on the BSID given to all ages encompassed by the BSID. This particular sample specifically includes 18-month-olds only; therefore, it is possible that the factors derived in our subscales specifically reflected this particular age group. As mentioned before, language abilities “burst” at age 18-months, which is why this sample could strongly reflect linguistic ability within the factors.

### *Predictive Ability of the 10-month Tasks to 18-month BSID-II*

Once factors were derived from the BSID-II data of the 18-month old subjects, it was of interest to examine the predictive ability of the tasks given at the 10-month visits to the BSID-II at 18-months. To do this, factor scores for the three factors were figured for each of the subjects. Then, a multiple regression technique was used to examine the degree to which the Willatts Problem Solving Task, the Immediate Recall and Delayed Recall portions of the Elicited Imitation Task (both administered at 10 months), and the MacArthur-Bates Communicative Development Inventory (administered at 18 months) were associated with the factors. We also examined the association between these tasks and the MDI Standard Score of the BSID-II.

*Predictive Properties of Willatt's Problem-Solving Task.* The Willatts Problem-Solving Task was not correlated with the Bayley MDI composite. When subjects' factor scores were used as the dependent variables in this regression analysis, the Willatts Problem Solving Task was unrelated to the two language factors, but it did significantly predict to the Object Manipulation factor; however this relationship was interesting in that a negative correlation was found between The Willatts Problem Solving Task and Object Manipulation. One could speculate that if a 10-month-old infant scores high in problem solving, it is possible that by 18-months of age, these infants have mastered the items that appear on the BSID-II so well as to be bored by their administration; in other words, those 18-month-olds who did so well with problem solving at 10-months of age may have found the BSID-II items to be too easy for them. It will be of interest to see how these infants perform on later tasks of problem-solving done at later ages, as we have been following most of these subjects in a follow-up study throughout their preschool years. The tasks of interest would be the Tower of Hanoi Task, given at 4-years and 5-years of age, and the object manipulation portions of the Wechsler Preschool and Primary Scale of Intelligence, or WPPSI.

*Predictive Properties of the Elicited Imitation Task.* The Elicited Imitation Scores for immediate and delayed recall were unrelated to the BSID-II MDI composite scores and to any of the three factors derived from the BSID-II.

*MBCDI-WS.* After finding two of the factors to be strongly linked to language, we were interested to explore if these two factors were related to the number of words each subject was saying, as well as the average length of sentences each subject is producing at 18-months of age. The measure of sentence length from the MacArthur-Bates CDI-WS, was significantly correlated with both the Expressive Language factor ( $r=.440, p=.001$ ) and the Receptive Language factor ( $r= .400, p= .003$ ). The two factors did correlate with the MDI composite score but the correlation between the MBCDI and the factors was appreciably stronger, representing an eightfold increase in the proportion of variance accounted for (see *Table 8*). This indicates that a child's understanding of words at 18-months of age could be related to their performance on the BSID-II items that involve the skill of language comprehension. Second, the length of sentences that the child was producing by 18-months of age was reasonably related to a child's expressive language abilities, represented by Factor 1. Both of these relationships between the two tasks demonstrates predictability between two standardized measures (a developmental test and a language inventory), and also helps support that the MacArthur-Bates CDI-WS, a parent report item, may be the most accurate measure this research group can use to give an accurate account of a child's linguistic ability at 18-months of age.

Also as we expected, the information from the MBCDI was not correlated significantly with Object Manipulation (*Table 9*). In fact, the correlation was negative. This can be interpreted as a demonstration of how verbal ability and object manipulation are separate entities on the BSID-II, and a better understanding of a child's specific abilities does occur when specific factors are extracted from the BSID-II; language factors do not have an impact on object manipulation factors, or vice versa.

*Can specific measures predict global measures?*

The overall results from all of these analyses help us to answer the key question of this paper: Can specific measures of early cognition predict a later performance on a global measure such as a standardized developmental test? From this particular analysis, it appears that specific measures of cognitive ability are generally unrelated to how well a child will perform on a global measure such as the BSID-II later on.

The Willatts Problem Solving Task did, however, significantly predict to the factor specifically involved with skill in object manipulation. The immediate recall and delayed recall portions of the Elicited Imitation Task were not associated with any of the three factors. One explanation of these outcomes could be that the Willatts Problem Solving Task and Object Manipulation do not require language proficiency to complete, which may not be particularly strong at 18-months of age. Object Manipulation is most likely an ability that an infant has been using regularly since even before 10-months of age, and has improved gradually until the BSID-II was administered to the same infant at 18-months of age. As stated earlier, there could be a number of behavioral reasons as to why 10-month-old infants are successful at problem solving on the Willatts task perform poorly on Object Manipulation on the BSID-II. It was of interest to be sure that this was not due to statistical “artifacts” on this particular data set that could arise from using tetrachoric correlations. This particular type of correlation often creates negative relationships between items, which in turn, could have a negative effect on a child’s factor score if they successfully completed the item. It was verified that all BSID-II items that loaded onto Object Manipulation did only have a positive effect on a child’s factor score; therefore, successful completion of items in this factor did not cause any unwanted negative outcomes in the child’s factor score.

The fact that Elicited Imitation does not significantly demonstrate predictability to the three factor scores is mildly surprising. While the task does not involve linguistic ability per se, it does require an understanding of language to execute the steps; in addition, it clearly involves

manipulation of objects to reach the end goal. It could be that the Object Manipulation factor reflects an underlying ability of manipulating objects to plan and solve a means-end problem to reach a specific goal, not object manipulation required to imitate from memory.

This finding may lead one to wonder why some tests that focus on similar cognitive abilities seem to have little relationship with one another. For one thing, the BSID-II is a test of overall development of a child, and many cognitive abilities are required by the subject to complete the items on the BSID-II. Specific measures of cognition, including the Willatts Problem-Solving Task and Elicited Imitation Task, are tasks that utilize particular skills and abilities.

In this analysis, we were looking at predictability from one specific age to another specific age. The scales created by Kohen-Raz (1967) may have demonstrated better predictive ability because this scale was developed to include BSID performance of all age groups, and not as specific as the current analysis. It is possible that cognitive tasks given to infants at ages other than 10-months should be examined; there may be a particular age that offers better insight to how a child may perform on developmental outcomes at 18-months of age. Or, on the same token, a predicted age other than 18-months of age may be more reliable as a developmental outcome. It would also be interesting to see if Object Manipulation demonstrated predictability to other measures of problem-solving across different ages, and if it offered better predictability within the realm of problem-solving rather than other specific measures of cognition.

Overall, specific measures may be good indications of a child's ability at a particular point in development, and give researchers a good indication of where a child is in terms of cognitive ability at any given age. Global measures are an effective tool to see how a child is performing in their overall development. Because they are standardized, global measures of development are also useful in comparing individual differences in children of the same age group. According to this analysis, it seems that if a specific age is being examined, using a

factor analysis to derive factor scores may not be necessary if the outcome is a language component of a standardized developmental measure, but it may be promising if the ability that is being tested is planning or problem-solving. It is difficult to say whether one type of developmental measure is better than the other from this analysis because it is of importance to be able to: 1) make observations of specific cognitive abilities at specific ages, and 2) to have a “score” from an overall measure of development to have as a reference in comparisons of individual differences within children of a particular age group. Based on the results of this analysis, it is reasonable to say that predictability of specific measures of cognition, particularly problem solving, to global measures of development is moderately useful when looking at aspects of executive function, but not necessarily when linguistic ability is a predicted outcome. The use of factor analysis methods should be further utilized in these comparisons because they may be more sensitive to detecting the underlying abilities of toddlers on standardized developmental tests. Further testing of infants and toddlers at various points in development would be needed to confirm this idea.

## References

- Baddeley, A. & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation* (47-89). New York: Academic Press.
- Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J.H. Flavell & L.D. Ross (Eds), *Cognitive Social Development: Frontiers and possible futures*. New York: Cambridge University, 225.
- Bauer, P. & Shore, C. (1987). Making a memorable event: Effects of familiarity and organization on young children's recall of action sequences. *Cognitive Development*, 2, 327-338.
- Bauer, P. (2004). Getting explicit memory off the ground: Steps toward construction of a neuron-developmental account of changes in the first two years of life. *Developmental Review*, 24, 347-373.
- Bauer, P. (2004). New developments in the study of infant memory. In D.M. Teti (Ed.), *Handbook of research methods in developmental psychology*. Oxford, England: Blackwell.
- Bayley, N. (1993). *Manual for the Bayley Scales of Infant Development, 2<sup>nd</sup> Edition*. San Antonio, TX: The Psychological Corporation.
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, 29, 109-160.
- Browne, M. W., Cudeck, R., Tateneni, K., & Mels G. (2004). *CEFA: Comprehensive exploratory factor analysis* [WWW document and computer program]. Retrieved from <http://quantrm2.psy.ohio-state.edu/browne/>.
- Casey, B., Giedd, J., & Thomas, K. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, 54, 241-257.

- Chen, Z. & Siegler, R. (2000). Across the great divide: Bridging the gap between understanding of toddlers' and older children's thinking. *Monographs of the Society for Research in Child Development*, 65, v-96.
- Colombo, J. & Cheatham, C.L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. In R. Kail (ed.), *Advances in Child Development and Behavior, Volume 34*, (pp. 283-310). Oxford: Academic Press.
- Dale, P., Bates, E., Reznick, J., & Morisset, C. (1989). The validity of a parent report instrument for child language at twenty months. *Journal of Child Language*, 16, 239-249.
- Kohen-Raz, R. (1967). Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley infant scale of mental development. *Genetic Psychology Monographs*, 76, 3-21.
- Diamond (1988). The abilities and neural mechanisms underlying A-not-B performance. *Child Development*, 59, 523-527.
- Diamond (2006). The early development of executive functions. In E. Bialystok & I.M. Craik (Eds.), *Lifespan Cognition: Mechanisms for Change* (pp. 70-95). New York, New York: Oxford University Press, Inc.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates E., Thal, D., Hartung, J. (1993). *Guide and technical manual for the MacArthur Communicative Development Inventories*. San Diego, CA: Singular Press.
- Hack, M., Taylor, H., Drotar, D., Schlucter, M., Carter, L., Wilson-Costello, D., Klein, N., Friedman, H., Mercuri-Minich, N., & Morrow, M. (2005). Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. *Pediatrics*, 116, 333-341.
- Huttenlocher, P. (1994). Synaptogenesis, synapse elimination, and neural plasticity in human cerebral cortex. In C.A. Nelson (Ed.) *Threats to optimal development*. The Minnesota Symposium on Child Psychology, Vol. 27, 35-54. Hillsdale, NJ: Erlbaum.

- Kubinger, K. (2003). On artificial results due to using factor analysis for dichotomous variables. *Psychology Science*, 45, 106-110.
- Lichtenberger, E. (2005). General measures of cognition for the preschool child. *Mental Retardation and Developmental Disabilities Research Reviews*, 11, 197-208.
- Meltzoff, A. (1985). Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Development*, 59, 62-72.
- Meltzoff, A. (1988). Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development*, 59, 217-225.
- Messer, D., McCarthy, M., McQuiston, S., MacTurk, R., Yarrow, L., & Vietze, P. (1986). Relation between mastery behavior in infancy and competence in early childhood. *Developmental Psychology*, 22, 366-372.
- Molfese, V. & Acheson, S. (1997). Infant preschool mental and verbal abilities: How are infant scores related to preschool scores? *International Journal of Behavioral Development*, 20, 595-607.
- Parry, C. & McArdle, J. (1991). An applied comparison of methods for least-squares factor analysis of dichotomous variables. *Applied Psychological Measurement*, 15, 35-46
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Powell, K. & Voeller, K. (2004). Prefrontal executive function syndromes in children. *Journal of Child Neurology*, 19, 785-797.
- Rose, S., Feldman, J., & Wallace, I. (1992). Infant information processing in relation to six-year cognitive outcomes. *Child Development*, 63, 1126-1141.
- Rovee-Collier, C. (1997). Dissociations of infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, 104, 467-498.

Weibe, S. (2004). Tracing the development of executive function in the transition from infancy to early childhood: Relations between working memory, sustained attention, and inhibitory control. Ph.D. dissertation, University of Minnesota, United States -- Minnesota. Retrieved January 22, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 3107944).

Willatts, P. (1984b). The Stage-IV infant's solution of problem-solving requiring the use of supports. *Infant Behavior and Development*, 7, 125-134.

Willatts, P. (1999). Development of means-end behavior in young infants: Pulling a support to retrieve a distant object. *Developmental Psychology*, 35, 651-667.

## APPENDIX A

### Tables

**Table 1***Scoring Procedure for the Wilatts Problem-Solving Task*

---

Action of Infant	Trial 1	Trial 2	Trial 3
Removal of Barrier	0 1 2	0 1 2	0 1 2
Look at Checkered Cloth	0 1 2	0 1 2	0 1 2
Grasps Checkered Cloth	0 1 2	0 1 2	0 1 2
Pull Checkered Cloth	0 1 2	0 1 2	0 1 2
Look at Black Cover Cloth	0 1 2	0 1 2	0 1 2
Grasps Black Cover Cloth	0 1 2	0 1 2	0 1 2
Removes Black Cover Cloth	0 1 2	0 1 2	0 1 2
Looks at Toy	0 1 2	0 1 2	0 1 2
Retrieves Toy	0 1 2	0 1 2	0 1 2
Total			AVG=

*For each action taken by the infant, a score of 0, 1, or 2 is given. A score of 0 = no action, 1= action is completed, but without clear intention, and 2= action is completed with clear intention. These scores are then totaled for each of the three trials and an average score is calculated from these three trials.*

**Table 2***Descriptive Statistics for 10-Month and 18-Month Subjects*

Task	Mean	SD	Ranges
BSID-II MDI Raw Score	111*	5.00*	102-125
BSID-II MDI Standardized Score	98*	9.00*	79-125
Willatts Problem Solving Task	6.6	5.60	0-17.7
Elicited Imitation Immediate Recall	1.04	0.42	0-2
Elicited Imitation Delayed Recall	0.70	0.49	0-2
MBCDI – Sentence Length	1.69	.99	1-4
MBCDI – Word Count	73.16	83.08	2-385

Information for the study sample for tasks completed at the 10-month visit and the 18-month visit (n=104).

\*In the BSID-II Test Manual, it is stated that the MDI was piloted in a sample (n=200), and the average raw score was 112.4, SD = 7.5. The average standardized score was 99.8, SD=15 (Bayley 1993). No averages are available for the Willatts Problem Solving Task or the Elicited Imitation Task.

**Table 3**

*Factor Correlations*

---

	Factor 1	Factor 2	Factor 3
Factor 1	1.000		
Factor 2	0.244	1.000	
Factor 3	0.011	0.106	1.000

---

*The Three Factors:* 1) Expressive Language, 2) Receptive Language, and 3) Object Manipulation, are not significantly correlated with each other.

**Table 4***Factors extracted from BSID-II MDI subjects' item responses.*

<i>Factor</i>	<i>Item #</i>	<i>Description</i>	<i>Loadings</i>
Expressive Language	106	Words to Make Wants Known	.749
	109	Names One Picture	.573
	113	Says Eight Words	.803
	114	Uses a Two-Word Utterance	.814
	117	Imitates a Two-Word Sentence	.592
	121	Uses Pronouns	.863
	127	Uses a Three-Word Sentence	.970
Receptive Language	99	Points to Two Pictures	.470
	101	Shows Shoes, Clothing, or Other Object	.602
	108	Points to Three Doll Parts	.634
	109	Names One Picture	.532
	110	Names One Object	.530
	111	Combines a Word and a Gesture	.539
	118	Identifies Objects in Photograph	.701
	122	Points to Five Pictures	.460
	124	Discriminates Book, Cube, and Key	.597
	125	Matches Pictures	.848
126	Names Three Objects	.751	
Object Manipulation	105	Retrieves Toy (Clear Box)	.487
	112	Places Four Pieces (Blue Puzzle Board)	.911
	115	Completes Pink Board	.726
	120	Completes Reverse Pink Board	.803
	123	Builds a Tower of Six Cubes	.451

BSID-II MDI Items that load together into common factors. Three factors were extracted using OLS method of factor analysis. The factors were then rotated using an orthogonal-varimax rotation to increase interpretability.

**Table 5**  
*Loadings of Individual BSID-II Test Items onto Each Factor*

MDI Item	Description	Factor 1	Factor 2	Factor 3
97	Builds Tower of Two Cubes	-0.130	-0.003	0.352
98	Places Pegs in 70 seconds	0.047	-0.016	0.093
99	Points to Two Pictures	0.313	0.470	0.020
100	Uses Two Different Words Appropriately	0.317	0.278	0.316
101	Shows Shoe, Other Clothing, or Object	-0.071	0.602	-0.088
102	Retrieves Toy (Visible Displacement)	-0.168	0.344	0.361
103	Imitates Crayon Stroke	-0.311	0.112	0.050
104	Uses Rod to Attain Toy	-0.254	-0.153	-0.016
105	Retrieves Toy (Visible Displacements)	0.233	0.174	0.487
106	Uses Words to Make Wants Known	0.749	0.367	0.009
107	Follows Direction (Doll)	0.039	0.144	0.107
108	Points to Three of Doll's Body Parts	0.263	0.634	-0.166
109	Names One Picture	0.573	0.532	-0.063
110	Names One Object	0.379	0.530	0.116
111	Combines Word and Gesture	0.382	0.539	0.219
112	Places Four Pieces in 150 seconds	0.006	0.169	0.911
113	Says Eight Different Words	0.803	0.479	-0.008
114	Uses a Two-Word Utterance	0.814	0.206	0.104
115	Completes Pink Board	-0.083	-0.087	0.726
116	Differentiates Scribble from Stroke	-0.294	0.079	-0.048
117	Imitates Two-Word Sentence	0.592	0.450	-0.211
118	Identifies Objects in Photographs	0.140	0.701	-0.097
119	Places Pegs in 25 seconds	0.292	0.215	-0.015
120	Completes Reversed Pink Board	0.014	-0.256	0.803
121	Uses Pronoun(s)	0.863	-0.108	0.055
122	Points to Five Pictures	0.169	0.460	0.193
123	Builds a Tower of Six Cubes	0.069	0.101	0.451
124	Discriminates Book, Cube, and Key	-0.125	0.597	-0.389
125	Matches Pictures	0.023	0.848	0.165
126	Names Three Objects	0.038	0.751	0.029
127	Uses a Three Word Sentence	0.970	-0.220	-0.107

*A summary of how each factor loaded onto each individual test item of the MDI portion of the BSID-II after the orthogonal varimax rotation was applied. These loadings were applied to each individual subjects' overall standardized score to assign subjects a factor score for each of the three factors.*

## Table 6

### *Formula for Estimating Factor Scores for Each Subject*

---

$$F_{\text{subject}} = a_1X_1 + a_2X_2 \dots a_nX_n$$

F = Factor Score

X = Each subject's individual response to a BSID-II item (1=Correct; 0=Incorrect)

a = Factor Loading associated with the specific factor in question

---

*Each subject (n=104) was assigned a factor score for each of the three factors based on the regression method for finding factor scores. Subjects received a "1" for a correct response on a BSID-II item, and a "0" for an incorrect response. The high loadings for each factor that corresponded with the rotated factor analysis were used to figure the factor score.*

**Table 7**

---

*Prediction of 10-month Tasks to 18-month BSID-II MDI Standard Scores*

---

 $R^2 = .040, F(1,103) = 1.404, p = ns$ 

	$\beta$	$t$	$P$
Willatts Problem Solving Task	0.133	1.359	0.177
Elicited Imitation, Immediate Recall	-0.076	-0.761	0.449
Elicited Imitation, Delayed Recall	0.147	1.467	0.145

---

A multiple regression technique was used in this analysis to demonstrate the predictability of tasks given to subjects at 10-months for age to the BSID-II Standard Score of each subject. The Willatts Problem-Solving Task and the Delayed Recall portion of the Elicited Imitation Task demonstrate moderate predictability, while the Immediate Recall Portion of the Elicited Imitation Task did not.

**Table 8**

	MDI Composite	Expressive Language	Receptive Language	Object Manipulation
Willatts	.133	.087	.001	-.192*
EI immediate	-.042	.089	.120	.014
EI delayed	.133	.076	.064	.031
MBCDI word count	.216*	.188	.400***	-.200
MBCDI sentence	.337*	.420***	.021	-.037

Using a correlational technique, it was found that the 10-month tasks predict identically to the standardized MDI scores of the BSID-II for each subject as each of the three factor scores derived for each subject.

\*\*\*  $p = .001$

\*  $p = .05$

**Table 9**

*Correlations between each of the predictor variables*

	WPST	EI: IR	EI: DR	MDI
Willatts Problem-Solving Task (WPST)	--	.027	.012	.133
Elicited Imitation: Immediate Recall (EI: IR)		--	.207	-.042
Elicited Imitation: Delayed Recall (EI:DR)			--	.133

Using a correlational method, above are the relationships between each of the tasks given to the subjects at 10-months of age and their BSID-II Mental Index Score. Note there are no strong relationships between any of the task