Behavioral characteristics during developmental assessments of preschool children with and without a maternal psychiatric history

Hannah J. Robinson

University of North Carolina at Chapel Hill

Spring 2015

A thesis presented to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Bachelor of Arts degree with Honors in Psychology.

Advisor _____
Dr. Barbara Davis Goldman

Reader _____
Dr. F. Charles Wiss

Reader _____
Rebecca Stephens

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Barbara Davis Goldman, for the continuous support of my research, for her guidance, enthusiasm, and invaluable knowledge. She has shaped my intellectual development by encouraging me to think creatively about research questions, data, and methodology. Her willingness to invest herself into the guidance and support of young researchers and her ability to encourage the pursuit of learning and love for your work is unrivaled.

Secondly, I would like to thank the other members of my committee and faculty supporters to whom I am deeply indebted. I am grateful to Dr. Charlie Wiss and Rebecca Stephens, for their time, expertise, and willingness to edit my thesis and provide valuable suggestions and comments. I would like to thank Dr. Peter Ornstein for his enthusiasm, help in conducting analyses, and the presentation process. I would also like to thank Dr. Brad Hammer for providing me with the guidance and direction that led me to pursue research and an honors project.

Thirdly, I would like to thank Maggie Fox and Mallory Turner, lab staff, for their valuable assistance in preparing the datasets that I used and their many contributions to my thesis work. They have provided me with their crucial feedback and encouragement, as well as joyful conversations during my time in the lab. I would like also like to thank Adeline Brooks, Lauren Conder, Maggie Fox, Gabby McHarg, Angelique Ray, and Mallory Turner for their highly reliable original rating data, which I used to answer my own questions.

Finally, I would like to thank my friends and family for their support. I am grateful to Allie Pinosky, Delaney McGuirt, Ian Hawkins Katie Martinez, Morgan May, Sinclair McLean and the Hawkins family for cheering me on every step of the way.

Abstract

Objective: The primary aim of this study was to examine behavioral differences at age four years between child offspring of mothers with a psychiatric diagnosis that included psychosis and child offspring of mothers who have never received such a diagnosis. *Method*: This study involved secondary data analysis of child behavior ratings made for quality control purposes from observations of video records of individual developmental assessments. In this thesis, these ratings are used to characterize the children's behavior directly so as to compare groups of children with and without possible increased risk of psychiatric disorders due to genetic history. As part of this thesis, inter-rater reliability across multiple raters was also established. *Results:* Independent-samples *t*-tests were used to compare the data between the two groups on overall mean ratings, on the seven behavioral dimensions, and the overall average behavior rating. Compared to the comparison group (C), children at genetic risk for developing a psychiatric disorder (R) exhibited significantly lower levels of appropriate activity level, focus, cooperativeness, and persistence during the testing session. Males at genetic risk for developing a psychiatric disorder exhibited lower levels of all behavioral dimensions when compared to male comparisons. Similarly, females at genetic risk for developing a psychiatric disorder exhibited lower levels of appropriate activity level, focus, cooperativeness, persistence, and enthusiasm compared to comparison females. *Limitations*: The context for the ratings may not represent the child's typical behavior in other contexts. Due to the limited sample size, data from offspring of mothers with different psychiatric diagnoses were grouped together, and were not examined separately. *Conclusions:* Assessment of behavior characteristics may be useful in characterizing children at genetic risk for developing a psychiatric disorder.

Behavioral characteristics during developmental assessments of preschool children with and without a maternal psychiatric history

Mental illness is recognized as a leading cause of morbidity worldwide (Gillam, Yates, & Badrinath, 2012). In the United States alone, an estimated 26.2 percent of Americans ages 18 and older suffer from a diagnosable mental disorder (Kessler, Chiu, Demler, & Walters, 2005). Genetic factors are known to play an important role in influencing susceptibility to many mental disorders, including schizophrenia, schizoaffective disorder, and bipolar disorder. Biological offspring of parents with these disorders are at an increased risk of developing a neurodevelopmental or psychiatric disorder, including developing the same disorder and/or related psychopathology later in life. Schizophrenia, schizoaffective disorder, and bipolar disorder, for example, develop through a complex interplay of genetic susceptibility and other risk factors; however, the etiology of these disorders still remains poorly understood. Through previous research, temperament traits have been demonstrated to predict (Chang, Blasey, Ketter, & Steiner, 2003) and predispose (Akiskal & Akiskal, 2005) individuals to developing various psychiatric disorders. Thus, examining behavioral characteristics that contribute to the formation of temperament in offspring of parents with schizophrenia or bipolar disorder, for example, may be useful. Such examination might be able to provide specific characteristics that serve as potential vulnerability indicators and targets for intervention programs for children at risk for developing these and other disorders. The current study is a secondary data analysis of ratings of child behaviors obtained by multiple, trained raters during developmental assessments at age four. The intent is to examine behavioral differences between children of mothers who had received psychiatric diagnoses of psychosis (in this sample, primarily schizophrenia and

schizoaffective disorder, and bipolar disorder, all with psychosis) and children of mothers who had never received a psychiatric diagnosis.

The high prevalence and morbidity of serious psychiatric conditions, such as schizophrenia, schizoaffective disorder, and bipolar disorder underscores the importance of understanding the risk factors and vulnerability indicators involved to improve the early identification and treatment of these disorders and related psychopathology. Schizophrenia affects 1.1% of the United States adult population, affecting males and females equally (Kessler et al., 2005). Psychotic features, such as delusions and hallucinations, poor cognitive functioning, and dysfunctional or disorganized ways of thinking, characterize schizophrenia (National Institute of Mental Health [NIMH], 2009). Because the causes of schizophrenia are still relatively unknown, treatments, such as antipsychotic medications and psychosocial treatments for the disorder focus on minimizing or eliminating the symptoms rather than focusing directly on the disorder itself. This requires long-term medical and social care for those with the disorder (National Institute of Mental Health [NIMH], 2009).

Bipolar disorder is a severe mood disorder affecting 2.6% of the United States adult population (Kessler et al., 2005). This disorder occurs in both males and females. Bipolar disorder is characterized by intense emotional states that range from mania to severe depressive episodes. Psychotic features and cognitive disturbances often accompany these emotional states. Treatment of the disorder usually involves medication, including mood stabilizers, atypical antipsychotics, and antidepressants, in combination with psychotherapy that focuses on lifestyle management. Proper treatment helps individuals gain better control over mood swings and other symptoms (National Institute of Mental Health [NIMH], 2012).

The high prevalence of individuals diagnosed with psychiatric disorders is an important consideration given that two-thirds of these women and over half of these men are parents (Kessler et al., 2005). The offspring of parents with a severe mental disorder (SMD) are at an increased risk of developing a range of mental disorders themselves, particularly offspring with two affected parents (Dean et al., 2010). Even though it is extremely difficult to determine how much of the impact of parental mental illness stems from genetics, and how much stems from the environment of living with, or being separated from, a parent with mental illness, given the evidence, understanding the risk factors as well as related protective factors that give rise to familial vulnerability to mental disorders is important to support this population of offspring of parents with severe mental illness.

Understanding a child's temperament and its influence on the way he or she interacts with others, as well as behaviors involved in the development of temperament, can help clinicians and families better understand how a child relates to and reacts to the environment around the child. This information can guide caregivers to support children in the development of strengths that he or she may need to succeed in future relationships and new environments.

Although no single definition of temperament has gained universal approval among researchers, most agree that temperament involves individual emotional and behavioral differences in reactivity, mood, and sensitivity. Based on these traits, some researchers commonly categorize children into three temperament types: easy/flexible, active/feisty, and slow to warm up/cautious. Easy/flexible children tend to be joyful, adaptable, relaxed, and not easily angered or frustrated. They are also characterized as having regular sleeping and eating habits, which also contributes to their being perceived as "easy" because of their predictability. Active/feisty children may be easily upset by noise and other stimuli, intense in their reactions,

and fearful of new people or situations. These children may have irregular sleeping and eating habits. Slow to warm up/cautious children may be fussy or withdrawn, and react negatively to new people or situations. However, with repeated exposure, they may become more calm and happy in new situations (Thomas, Chess, Birch, Hertzig, & Korn, 1963).

Origins of these temperamental differences, whether biological or due to environmental factors, are highly debated. In the last decade, a broad consensus has been reached among researchers establishing that temperament has a biological basis (Kagan, 2005). Family, twin/sibling, and adoption studies have highlighted the importance of genetics in the development of temperament because genetic factors may elicit different environments or they may mediate responses to different environments. Such studies find that monozygotic (identical) twins are more similar than dizygotic (fraternal) twins when comparing temperament characteristics. In a sample of 53 same-sexed twin pairs, Torgersen and Kringlen (1978) found that monozygotic twins were more similar than dizygotic twins in three temperamental categories at two months of age. At nine months, monozygotic twins were more similar than dizygotic twins for all temperamental variables that were measured: activity level, rhythmicity, approach/withdrawal, adaptability, intensity, threshold, mood, distractibility, and attention span/persistence (Torgersen & Kringlen, 1978).

Environmental influences also play crucial roles in the development of temperament and behavior. Aspects of the environment, such as parenting, have important long-term consequences for the development of a child. This is important because literature highlights the negative effects of mental illness on parenting (Oyserman, Mowbray, Meares, & Firminger, 2000). A study has shown that family environmental factors, like parenting style and parent-child interaction, might moderate the relationship between specific child temperamental characteristics and child

internalizing and externalizing behavior (Bates, Pettit, Dodge, & Ridge, 1998). As noted earlier, environmental influences can be difficult to separate from genetic influences on child behavior in general, but they can have considerable impact, especially in situations where parenting behavior may be inconsistent and hard for the child to understand.

"Temperament" differs from "behavior" in that temperament is measured as a trait of an individual, combining the array of behaviors that an individual expresses at a specific period in life, and/or over the course of a lifetime. In contrast, behaviors are reactions to specific situations and stimuli. These expressed behaviors can be thought of as the state that an individual is in during a given situation. Together, relatively consistent behaviors across various contexts can be viewed as an individual's temperament.

Temperament traits have been proposed to predict the development of psychopathology later in life. For example, arrhythmicity, inflexibility, and high distractibility were associated with increased substance use, depressive symptoms, and delinquent activity in a sample of 297 high school students (Windle, 1991). Similarly, increased activity levels were associated with increased substance use and decreased mood in a sample of adolescents (Wills et al., 1995). In another study, children identified as 'behaviorally inhibited' at 21 months of age were found to have higher levels of anxiety disorders at age 7.5 than was the group earlier identified as 'behaviorally uninhibited' (Biederman et al., 1990).

Recent findings suggest that offspring of parents with bipolar disorder exhibit distinct temperamental traits as compared to healthy offspring of psychiatrically well parents (Singh, DelBello, & Strakowski, 2008). Specifically, offspring of parents with bipolar disorder exhibited a decreased tendency to approach new situations, flexibility, ability to follow the same daily sleep patterns, and persistence on tasks as compared to the comparison offspring group. In

addition, the results supported data from prior studies that demonstrated a predictive relationship between temperament and psychopathology (Chang et al., 2003). Chang, Blasey, Ketter, and Steiner (2003) found that bipolar offspring with psychiatric disorders had decreased flexibility, decreased ability to demonstrate persistence on tasks without distraction, and negative moods as compared to bipolar offspring without psychopathology. However, this particular study was limited by the lack of a comparison group of healthy offspring to compare to the bipolar offspring sample. The findings of Singh, DelBello, Strakowski (2008) and Chang et al. (2003) are also consistent with a study conducted in 2007 that did compare offspring of parents with bipolar disorder to offspring of healthy parents (Duffy et al., 2007). Duffy et al. (2007) confirmed a positive correlation between temperament and lifetime psychopathology. More specifically, emotionality was positively associated with the risk of having a mood disorder (Duffy et al., 2007). Emotionality can be described as high emotions and a tendency to cry easily and react intensely when upset, as defined by the self-reported Emotionality Activity Sociability (EAS) Temperament Survey for Children (Buss and Plomin, 1984).

Research on child temperament has identified a number of behavioral dimensions that can be used to measure temperament. Through the use of interview data of 22 parents of infants two to six months from the New York Longitudinal Study (NYLS), Thomas, Chess, Birch, Hertzig, and Korn (1963) identified nine behavioral categories that define temperament in infancy and early childhood: Activity Level, Rhythmicity, Distractibility, Approach versus Withdrawal, Adaptability, Attention Span and Persistence, Intensity of Reaction, Threshold of Responsiveness, and Quality of Mood. These nine dimensions have been widely used in research on child temperament. However, the dimensions have undergone change due to questions about the conceptual overlap of the scales.

Since then, much of the work on temperament in the 1970s and 1980s involved the development of temperament measures. The components of temperament scales are designed to be conceptually independent of one another, unlike the nine dimensions defined by Thomas et al. (1963). Many questionnaires have been created to assess temperament in young children based on previous research on temperament:

> The most widely used infant and childhood temperament questionnaires in the behavior—genetic literature have been the series of questionnaires designed by A. H. Buss, Plomin, Rowe, and colleagues: the Emotionality, Activity, Sociability, Impulsivity (EASI) Temperament Survey, the derived Emotionality, Activity, Sociability (EAS) questionnaire, and the Colorado Childhood Temperament Inventory (CCTI). The EASI and EAS questionnaires were developed in conjunction with A. H. Buss and Plomin's (1975) theory of temperament, whereas the CCTI was developed as a merger of the EASI dimensions and the New York Longitudinal Study (NYLS) approach. (Goldsmith, Lemery, Buss, & Campos, 1999, pp. 972).

In most recent years, the most widely used temperament questionnaire for young children has been Rothbart's Children's Behavior Questionnaire (CBQ) (Rothbart et al., 2001; Putnam & Rothbart, 2006).

In previous studies conducted in this field, parental or self-report questionnaires were used to assess temperament and psychopathology. Questionnaires have proven to be useful sources of information on the structure of temperament (Rothbart & Bates, 1998). However, controversy exists over the validity of parental descriptions of behavior because the correlation between the parental reports, usually based on questionnaires, and direct behavioral observations is usually low, with the correlations being less than 0.3 (Seifer et al., 1994; Biship, Spence, & McDonald, 2003). Of course, it is difficult to determine the source of the difference, given that multiple features, who provides the information, the context of the data collection, the method, and the duration/extent of the database, that is, a single observation vs. life with a child. are quite different. However, observational measures of behavior, relating to temperament, have been

relatively neglected in the research. Despite that, the current study uses secondary data analysis of direct behavioral ratings of observed behavior in a controlled assessment context in a child laboratory setting to examine behavioral characteristics in children at age four.

The environmental setting of the child assessment laboratory, somewhat similar to school but quite different in that other children are not present and the tester is not familiar to the child, likely arouses distinct behavioral characteristics in four-year-olds because of its novelty. Most four-year-olds have not been introduced to a one-to-one, structured, yet reasonably playful environment, which requires them to answer questions and perform new tasks, mostly at a child-sized table, upon request, even if they have been attending preschool or group childcare. Importantly, however, because the children's ratings are obtained from essentially the same environmental context across all the children, this uniformity allows for comparisons between groups of children when looking at behavioral (state) differences.

The primary aim of the current study was to examine behavioral differences between children of mothers with a psychiatric diagnosis with psychosis, primarily schizophrenia, schizoaffective disorder, or bipolar disorder with psychotic features, and children of mothers without any psychiatric diagnosis. Thus, the study asks whether there will be behavioral differences between children in an assessment situation when being tested one-on-one with trained testers who were blind to maternal history. The behavioral scores used in this secondary data analysis were obtained from the ratings of trained raters who were also blind to condition in four-year-old children, made for the purpose of quality control. The data analyzed were drawn from extant data about comparison (C) children and children at some genetic risk (R) of future psychiatric and/or neurodevelopmental disorders. On the basis of prior research, it was

hypothesized that R children would exhibit distinct behavioral profiles as compared to the behavioral profiles of C (comparison) children. We approached this as an exploratory study.

Method

Participants

All behavioral ratings (the data for the current study) already existed as part of another study; this is primarily a secondary data analysis that also uses existing data to create a new summary variable for each data record, and that determined the inter-rater reliability of the ratings. There are no true "participants" in this study, but instead there are de-identified data records. The data records were given new, thesis-research-specific IDs; a gender code (male, female) were also included with the behavioral ratings from however many raters had observed and rated that particular child's assessment session. As noted earlier, the researcher conducting this work at no time was aware of the identity of the children who had been observed and rated. The director of the child assessment lab, also the faculty supervisor of this thesis, was the person who prepared the "coding" of the individual records as belonging to a **C**omparison child or a child at **R**isk for a later disorder due to maternal history.

Data records were selected by the assessment lab director based on the availability of assessment sessions at age four that had been rated by at least two trained raters for quality control purposes. The researcher who conducted this thesis research had performed those kinds of ratings as part of her PSYC 395 research internship in an earlier semester and as part of her continued training in understanding those behavioral ratings more recently. Data records selected to examine in this study included those of 100 comparison four-year olds (52 males, 48 females) and 34 "at genetic risk" four-year olds (22 males, 12 females). The total number of data records analyzed was 134.

Background information on the original sample: The participants were originally recruited for research studies that were focused on early brain and behavior development in children whose mothers had diagnoses of schizophrenia, schizoaffective disorder and bipolar disorder, as well as the occasional other psychiatric disorder such as major depression with psychosis and psychosis not otherwise specified, and children of mothers without such diagnoses, nor any other psychiatric diagnosis. None of the data analyzed in this study were from children who had received any kind of prenatal or antenatal diagnosis of a brain malformation or neurodevelopmental disorder.

Measures

Diagnostic assessments

Maternal diagnoses were confirmed from medical records by psychiatrists affiliated with the larger study. Only a single category, "genetic risk" (R), was used here. Neither the child testers, nor the raters who observed the videos of the assessment sessions, were aware of whether the child was one at risk or was a comparison child. The researcher conducting this secondary data analysis has been given her own, coded data set with R and C added, but she has been given no other information related to the data records other than gender of the children.

Developmental assessment observational ratings

The developmental assessment at age four included six Stanford-Binet tests and other non-standardized assessments. The ratings of child behavior for the purpose of this study were those from about one-half of the assessment session, and included only the six sections (tests) of the comprehensive Stanford-Binet Intelligence Scales: Fifth Edition (Roid, 2003). The six sections are described below:

Object Series/Matrices. These tasks include a block pattern completion and picture pattern completion task, where the child must choose the block or answer choice that comes next in a given pattern by touching the corresponding item. This test also serves as a "routing" test to determine the level at which other nonverbal tests are entered.

Vocabulary. These tasks ask the child to describe the meaning of specific words, such as "cup" and "puddle". Then, the child must describe the actions that are occurring in a photo that is presented. This task includes answers such as "cutting" or "running". This test also serves as a "routing" test to determine the level at which other verbal tests are entered.

Visual-Spatial Processing (nonverbal). These tasks include items where the child must copy a picture of a "shape person" by using individual shape pieces to construct the figure, and completion of a Form Board activity where the child uses a plastic form with spaces for a triangle, a square, and a circle

Working Memory (nonverbal). These tasks include a block-tapping task, where the child must touch a series of blocks in such a way that exactly repeats the sequence of blocks tapped by the tester.

Fluid Reasoning (verbal). This test includes a "silly sentence" game where the child must indicate what is silly or impossible about a given sentence, and another where the child must organize pictures into matches of three and explain the reasons for the match. For example, one match may include a picture of a pen, pencil, and marker. The child must indicate that the three pictures create a match because they are all something that we write with.

Working Memory (verbal). The last test includes repeating specific sentences after hearing the tester say the sentence. The sentences get longer and more complex as the task continues.

Procedure

The children participating in the original study completed developmentally appropriate assessments at various ages; only ratings on the assessments conducted at age four are used in these analyses. The developmental assessments were video recorded, to be analyzed in more detail at a later time. All assessments were conducted at the Frank Porter Graham Child Development Institute, part of the University of North Carolina, located in Chapel Hill, North Carolina.

Behavior was assessed by trained raters (lab staff, students enrolled in PSYC 395 for a research internship, and student volunteers also involved in a research internship) using video recordings of the assessment at age four. Behavioral rating scores were obtained by raters viewing and rating observed behavior during the six portions of the Stanford-Binet described in the *Measures* section. The seven behavioral dimensions rated during each of the six sections were Alertness, Appropriate Activity Level, Focus, Freedom from Apprehension, Flexible Cooperativeness with Adult, Persistence, and Enthusiasm. Descriptions of these behavioral dimensions can be found in Table 1 of *Appendix A*. Behavioral ratings of each dimension ranged from 1 to 4, based on how frequently the behavior was observed, as follows: 1 (*low or absent*), 2 (*marginal/mixed*), 3 (*adequate*), and 4 (*clearly demonstrated*). Each of the six components of the Stanford-Binet, as well as each of the other parts of the assessment, not included here, were rated separately. The rater thus has to hold in mind for only a short period of time what has been observed so that a summary/averaged rating can be given. In addition, the ratings cover the entire assessment segment by segment because the original purpose of these ratings was for quality control—to determine if challenging behavioral states may have interfered with the validity of the scores obtained by the child. Child behavior, such as not being alert, or not focused, might

only interfere with validity on some but not all parts of the assessment. The child's behavior during the Stanford Binet was rated using a Likert scale on each of the seven behavioral dimensions plus an averaged-across-dimensions rating for that section, a total of six times, one rating per section, totaling 48 individual behavior ratings.

Assessment lab staff calculated an averaged rating for each of the seven behavioral dimensions (as a summary variable for each behavioral dimension across the six different Stanford-Binet tests) by averaging across the six rating scores from each section, for each rater.

Lastly, an overall average behavioral score was calculated by averaging the seven averaged scores obtained from averaging the six rating scores within each of the seven behavioral dimensions. All of the data were recorded using a data collection form created for the purposes of this study. This form can be found in *Appendix B*.

Seven raters extracted the behavioral data from the video recordings of the developmental assessments used in this secondary analysis. A form-scoring manual had been developed, which can be found in *Appendix B*. Each of the seven raters had been trained by assessment lab staff in using this manual to ensure consistent scoring between raters. Before rating the videos that were used in this sample, raters practiced using the rating scale by scoring several "practice" assessments and comparing their scores with a master key. A master key of one of these practice videos can be found in *Appendix B*.

The majority of data records in the sample used for this secondary data analysis were comprised of ratings that were averaged across two raters. Only 26 of the 134 ratings were from only one rater. Inter-rater reliability was assessed for each of the seven behavioral dimensions, across multiple combinations of raters. Eighty percent (n = 108) of the assessments were rated twice by two different raters. Inter-rater reliability was assessed separately for each of the seven

behavioral dimensions. Difference scores were calculated for each behavior dimension to assess the average difference between the two rater's scores. A sample of the 16 unique rater dyads was selected to assess inter-rater reliability. The five dyads that rated the greatest number of assessments were chosen for the inter-rater reliability sample. All seven raters were represented. A graph of these differences scores can be found in Graph 1 of *Appendix A*.

Results

All statistical analyses were performed using SPSS.

Using an alpha level of .05, a two-tailed independent-samples *t* test was conducted to evaluate whether children at genetic risk for developing a neurodevelopmental or psychiatric disorder and comparison children differed significantly on the average behavior rating for each of the seven behavioral dimensions. The proportion of males ($n = 74$) to females ($n = 60$) in the entire sample was reasonably balanced, but that was not true for the proportions for the "at genetic risk" group. The proportion of males to females was 22:12 in the "at genetic risk" group and 52:48 in the comparison group. The gender distribution for the "at genetic risk" group was not comparable; therefore, a two-tailed independent-samples *t* test was conducted to examine groups by gender. Lastly, we conducted independent-samples *t* tests to examine the males at genetic risk with the male comparisons and the females at genetic risk with the female comparisons.

The test comparing the at genetic risk group (R) with the comparison group (C) was significant for four of the seven dimensions: activity level, $t(43.88) = -2.68$, $p = .010$, focus, $t(41.94) = -2.65$, $p = .011$, flexible cooperativeness, $t(40.71) = -2.546$, $p = .015$, and persistence, $t(41.26) = -2.26$, $p = .029$. An examination of the group means indicate that the C group ($M = 3.34$, $SD = 0.50$) scored significantly higher on the average behavior rating for activity level than

the R group ($M = 2.98$, $SD = 0.74$). The C group ($M = 3.44$, $SD = 0.42$) scored significantly higher on the average behavior rating for focus than the R group ($M = 3.11$, $SD = 0.68$). The C group ($M = 3.60$, $SD = 0.44$) scored significantly higher on the average behavior rating for flexible cooperativeness than the R group ($M = 3.25$, $SD = 0.76$). The C group ($M = 3.51$, $SD = 0.41$) scored significantly higher on the average behavior rating for persistence than the R group ($M = 3.23$, $SD = 0.70$).

In addition, there was a significant difference between the C group and the R group for the averaged behavior rating of the seven behavior dimensions, $t(42.13) = -2.43$, $p = .019$. An examination of the group means indicate that the C group ($M = 3.60$, $SD = 0.30$) scored significantly higher on the average behavior rating for the seven behavior dimensions than the R group ($M = 3.39$, $SD = 0.47$). Statistics for all behavior dimensions are summarized in Table 2 of *Appendix A*.

A two-tailed independent-samples *t* test was conducted to evaluate whether males and females differed significantly on the average behavior rating for each of the seven behavior dimensions. The test was significant for three of the seven dimensions: flexible cooperativeness, $t(130.27) = 2.46$, $p = .015$, persistence, $t(132) = 2.27$, $p = .025$, and enthusiasm, $t(126.879) = 3.288$, $p = .001$. An examination of the group means indicate that females ($M = 3.64$, $SD = 0.44$) scored significantly higher on the average behavior rating for flexible cooperativeness than males ($M = 3.41$, $SD = 0.62$). Females ($M = 3.55$, $SD = 0.44$) scored significantly higher on the average behavior rating for persistence than males ($M = 3.35$, $SD = 0.56$). Females ($M = 3.68$, $SD = 0.37$) scored significantly higher on the average behavior rating for enthusiasm than males ($M = 3.42$, $SD = 0.56$).

In addition, there was a significant difference between females and males for the average behavior rating of the seven behavior dimensions, $t(129.288) = 2.69$, $p = .008$. An examination of the group means indicate that females ($M = 3.63$, $SD = 0.28$) scored significantly higher on the average behavior rating for the seven behavior dimensions than males ($M = 3.48$, $SD = 0.40$). Statistics for all behavior dimensions are summarized in Table 3 of *Appendix A*.

A two-tailed independent-samples *t* test was conducted to evaluate whether comparison males and comparison females differed significantly on the average behavior rating for each of the seven behavior dimensions. The test was significant for two of the seven dimensions: flexible cooperativeness, $t(88.38) = 2.43$, $p = .022$ and enthusiasm, $t(86.50) = 2.97$, $p = .004$. An examination of the group means indicate that female comparisons ($M = 3.71$, $SD = 0.33$) scored significantly higher on the average behavior rating for flexible cooperativeness than male comparisons ($M = 3.51$, $SD = 0.50$). Female comparisons ($M = 3.71$, $SD = 0.34$) scored significantly higher on the average behavior rating for enthusiasm than male comparisons ($M = 3.45$, $SD = 0.54$).

In addition, there was a significant difference between female comparisons and male comparisons for the averaged behavior rating of the seven behavior dimensions, $t(92.78) = 2.08$, $p = .040$. An examination of the group means indicate that female comparisons ($M = 3.66$, $SD = 0.24$) scored significantly higher on the averaged behavior rating for the seven behavior dimensions than male comparisons ($M = 3.54$, $SD = 0.33$). Statistics for all behavior dimensions are summarized in Table 4 of *Appendix A*.

A two-tailed independent-samples *t* test was conducted to evaluate whether females at genetic risk for developing a psychiatric disorder and males at genetic risk for developing a psychiatric disorder differed significantly on the average behavior rating for each of the seven

behavior dimensions. The test was significant for only one of the seven dimensions: alertness, $t(23.90) = 2.20$, $p = .038$. An examination of the group means indicated that females at genetic risk ($M = 3.98$, $SD = 0.06$) scored significantly higher on the average behavior rating for alertness than males at genetic risk ($M = 3.82$, $SD = 0.31$). Statistics for all behavior dimensions are summarized in Table 5 of *Appendix A*.

A two-tailed independent-samples *t* test was conducted to evaluate whether females at genetic risk for developing a psychiatric disorder and female comparisons differed significantly on the averaged behavior rating for each of the seven behavior dimensions. The test was significant for only one of the seven dimensions: alertness, $t(57.34) = -2.28$, $p = .026$. An examination of the group means indicated that females at genetic risk ($M = 3.98$, $SD = 0.06$) scored significantly higher on the average behavior rating for alertness than female controls ($M = 3.89$, $SD = 0.23$). Statistics for all behavior dimensions are summarized in Table 6 of *Appendix A*.

Lastly, a two-tailed independent-samples *t* test was conducted to evaluate whether males at genetic risk for psychiatric disorders and male comparisons differed significantly on the averaged behavior rating for each of the seven behavior dimensions. The test was significant for three of the seven dimensions: appropriate activity level, $t(30.19) = 2.15$, $p = .040$, focus, $t(72) = 2.98$, $p = .004$, and persistence, $t(72) = 2.09$, $p = .040$. An examination of the group means indicate that male controls ($M = 3.29$, $SD = 0.55$) scored significantly higher on the average behavior rating for appropriate activity level than males at genetic risk for developing a psychiatric disorder ($M = 2.89$, $SD = 0.78$). Male controls ($M = 3.41$, $SD = 0.48$) scored significantly higher on the average behavior rating for focus than males ($M = 2.99$, $SD = 0.69$). Lastly, male controls ($M = 3.43$, $SD = 0.47$) scored significantly higher on the average behavior

rating for persistence than males at genetic risk for developing a psychiatric disorder (M = 3.15, SD = 0.69). Statistics for all behavior dimensions are summarized in Table 7 of *Appendix A*.

Discussion

The primary aim of this study was to examine behavioral differences at age four between child offspring of mothers with a psychiatric diagnosis that included psychosis and child offspring of mothers who have never received such a diagnosis, or any other psychiatric diagnosis. The hypothesis that children at genetic risk for developing a psychiatric disorder would exhibit distinct behavioral profiles as compared to the behavioral profiles of comparison children was supported by significant behavior differences across multiple behavior dimensions. Specifically, the R group exhibited differences in activity level, focus, flexible cooperativeness, and persistence compared to the C group. More specifically, the R group exhibited more inappropriate activity levels compared to the C group. These activity levels, including excessive fidgeting and squirming and unwillingness to sit in order to complete a given task, are less suited for the testing environment and similar situations. When compared to the C group, the R group also exhibited decreased levels of focus, cooperativeness, and persistence during the testing session, supporting data from another study examining temperament that showed that offspring of parents with bipolar disorder had a decreased tendency to approach new situations, decreased flexibility and decreased persistence on tasks as compared to the comparison offspring group (Singh, DelBello, & Strakowski, 2008). In contrast, in the present secondary data analysis, significant behavior differences did not exist for alertness, freedom from apprehension, and enthusiasm. Thus, some behavior dimensions may be better suited than others when used as an indicator of differences between these two populations. All group means can be found in Graph 2 of *Appendix A*.

Similar behavioral characteristic differences were found between the offspring with a maternal psychiatric history and the comparison offspring in the present study and other studies that examined measures of temperament. This may indicate that observed behavior at one time point can be used as an indicator of a child's temperament.

In order to examine possible gender differences in the behavior ratings, males and females were compared. Males exhibited significantly lower levels of cooperativeness, persistence, and enthusiasm when compared to females. Males were also much less focused (p = .051) when compared to females. Further comparisons were conducted to examine the effects of gender and whether or not the child had a maternal history of mental illness with psychosis or psychotic features. When compared to females, comparison males exhibited lower levels of cooperativeness, persistence (p = .055), and enthusiasm. However, males at genetic risk for developing a psychiatric disorder only scored significantly lower than females at genetic risk on alertness.

Males at genetic risk for developing a psychiatric disorder scored significantly lower on appropriate activity level, focus, and persistence when compared to comparison males. Importantly, males at genetic risk scored lower than male comparisons, female comparisons, and females at genetic risk on all seven different behavior dimensions. This could be due to gender differences in coping with a parental mental illness. For example, some studies have shown that boys are more adversely affected by a parent's depression. Boys were more likely to have poorer social competence and more behavior problems than girls (Gross, Conrad, Fogg, Willis, 1995).

Females at genetic risk for developing a psychiatric disorder scored significantly higher on alertness than female comparisons. Although not significantly different, females at genetic

risk also scored higher than female comparisons on the freedom from apprehension behavior dimension.

Given the results of this study, individual behaviors, as well as temperament due to is relationship to behavior, could serve as potential foci for the identification of, and possibly the appropriate intervention for, mental disorders in families with mothers who have a history of mental illness. The use of these targets for intervention could help delay or even prevent the onset, severity, or progression of mental disorders and related symptoms in children who have a family history of mental illness.

It is important to note the limitations of this study when interpreting the results. First, this study is limited by its small sample size of "at genetic risk" offspring. Due to this limited sample size, data records for offspring of specific psychiatric diagnoses were not examined separately. In future studies with larger samples, examination of offspring of mothers with different diagnoses may lead to different and more specific results. Separate analyses may also offer insight into behavior dimensions that can be targeted for intervention programs for offspring of specific disorders. In addition, the groups were not distributed evenly by gender. For example, the males at genetic risk (n = 22) in the sample almost double the number of females at genetic risk (n = 12) in the sample.

The study only examined ratings derived from observations of behavior rather than parent reports of behavior or self-reports of behavior. These ratings of fairly short segments of behavior may not be representative of the child. It is also important to consider that the children were in a novel environment where they were asked to answer questions and complete unfamiliar tasks. This environment may influence behaviors demonstrated during the assessment.

The children examined in this study were only four years of age. Although significant differences existed between the two groups, children may not exhibit exceptionally distinct behavioral profiles until much later on in life.

It is also important to note that not all of the children in the sample were living with their biological parents. Some of the children were adopted and others were being taken care of by other family members. Early childhood, including age four, may be a period in which offspring are particularly sensitive to parental mental illness because young children are completely dependent on their caregivers. The environmental factors associated with living with a mother or family member with a psychiatric history may play a role in the development of temperament and behavior.

Longitudinal studies that examine self- and parent-report data, as well as observational data, to determine if certain behaviors inform treatment response and prognosis in this population, are needed. Longitudinal studies would also be beneficial to discover when distinct behavioral profiles occur, or if they occur. In addition, further studies that assess the effectiveness of interventions that target these behavior characteristics in offspring of mothers with schizophrenia or bipolar disorder are necessary to support this vulnerable population.

References

Akiskal, K. K., & Akiskal, H. S. (2005). The theoretical underpinnings of affective temperaments: Implications for evolutionary foundations of bipolar disorder and human nature. *Journal of Affective Disorders*, *85*, 231-239.

Bates, J. E., Pettit, G. S., Dodge, K. A., & Ridge, B. (1998). Interaction of temperamental resistance to control and restrictive parenting in the development of externalizing behavior. *Developmental Psychology*, *34*(5), 982-995. doi: 10.1037/0012-1649.34.5.982

Biederman, J., Rosenbaum, J. F., Hirshfeld, D. R., Faraone, S. V., Bolduc, E. A., Gersten, M., Meminger, S. R., Kagan, J., Snidman, N., & Reznick, J. S. (1990). Psychiatric correlates of behavioral inhibition in young children of parents with and without psychiatric disorders. *Archives of General Psychiatry*, *47*, 21–26.

Biship, G. S., Spence, S. H., & McDonald, C. (2003). Can parents and teachers provide a reliable and valid report of behavioural inhibition? *Child Development*, *74*(6), 1899-1917.

Buss, A. & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, New Jersey: Erlbaum.

Chang, K. D., Blasey, C. M., Ketter, T. A., & Steiner, H. (2003). Temperament characteristics of child and adolescent bipolar offspring. *Journal of Affective Disorders*, *77*(1), 11-19. doi: 10.1016/S0165-0327(02)00105-2.

Dean, K., Stevens, H., Mortensen, P. B., Murray, R. M., Walsh, E., & Pedersen, C. B. (2010). Full spectrum of psychiatric outcomes among offspring with parental history of mental disorder. *Archives of General Psychiatry*, *67*, 822-829. doi: 10.1001/archgenpsychiatry.2010.86

Duffy, A., Alda, M., Trinneer, A., Demidenko, N., Grof, P., & Goodyer, I. M. (2007). Temperament, life events, and psychopathology among the offspring of bipolar parents [Abstract]. *European Child Adolescent Psychiatry, 16*(4), 222-228. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17136299

Gillam, S., Yates, J., & Badrinath, P. (Eds.). (2012). *Essential public health theory and practice, second edition*. New York, NY: Cambridge University Press.

Goldsmith, H. H., Lemery, K. S., Buss, K. A., & Campos, J. J. (1999). Genetic analyses of focal aspects of infant temperament. *Developmental Psychology*, *35*(4), 972-985. Retrieved from http://www.psych.stanford.edu/~carl/isl/PDFPublications/Genetic analyses of focal aspects of infant temperament.pdf

Gross, D., Conrad, B., Fogg, L., & Willis, L. (1995). A longitudinal study of maternal depression and preschool children's mental health. *Nursing Research, 44*(2), 96-101. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7892146

Kagan, J. (2005). Temperament. In R. E. Tremblay, R. G. Barr, R. Peters (Eds.), Encyclopedia on early childhood development (pp. 1-4). Montreal, Quebec: Centre of Excellence for Early Childhood Development.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*, *62*(6), 617-627.

Murray, C. J. L., & Lopez, A. D. (Eds.). (1996). *The global burden of disease: a comprehensive assessment of mortality, injuries, and risk factors in 1990 and projected to 2020*. Boston: Harvard School of Public Health.

Oyserman, D., Mowbray, C. T., Meares, P. A., & Firminger, K. B. (2000). Parenting among mothers with a serious mental illness. *American Journal of Orthopsychiatric, 70*(3), 296-315.

Putnam, S. P., & Rothbart, M. K. (2006). Development of Short and Very Short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment, 87*(1), 103-113.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales: Fifth Edition.* Itasca, IL: Riverside.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament as 3-7 years: The Children's Behavior Questionnaire. *Child Development*, *72*, 1394-1408.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. *Handbook of child psychology: Vol. 3. Social, emotional and personality development*. (5th ed.), 105-176. New York: Wiley.

Seifer, R. A., Sameroff, A. J., Barrette, L. C., Krafchuk, E. (1994). Infant temperament measured by multiple observations and mother report. *Child Development, 65*(5), 1478-1490.

Singh, M. K., DelBello, M. P., & Strakowski, S. M. (2008). Temperament in child offspring of parents with bipolar disorder. *Journal of Child and Adolescent Psychopharmacology, 18*(6), 589-593. doi: 10.1089/cap.2007.142

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.

Torgersen, A. M., & Kringlen, E. (1978). Genetic aspects of temperamental differences in infants: A study of same-sexed twins [Abstract]. *Journal of the American Academy of Child Psychiatry, 17*(3), 433-444. doi: 10.1016/S0002-7138(09)62299-8

U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Mental Health. (2009). *Schizophrenia*. (NIH Publication No. 09-3517). Retrieved from

http://www.nimh.nih.gov/health/publications/schizophrenia/index.shtml

U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Mental Health. (2012). *Bipolar disorder in adults*. (NIH Publication No. 02-2650). Retrieved from http://www.nimh.nih.gov/health/publications/bipolar-disorder-in-adults/index.shtml?rf#pub1

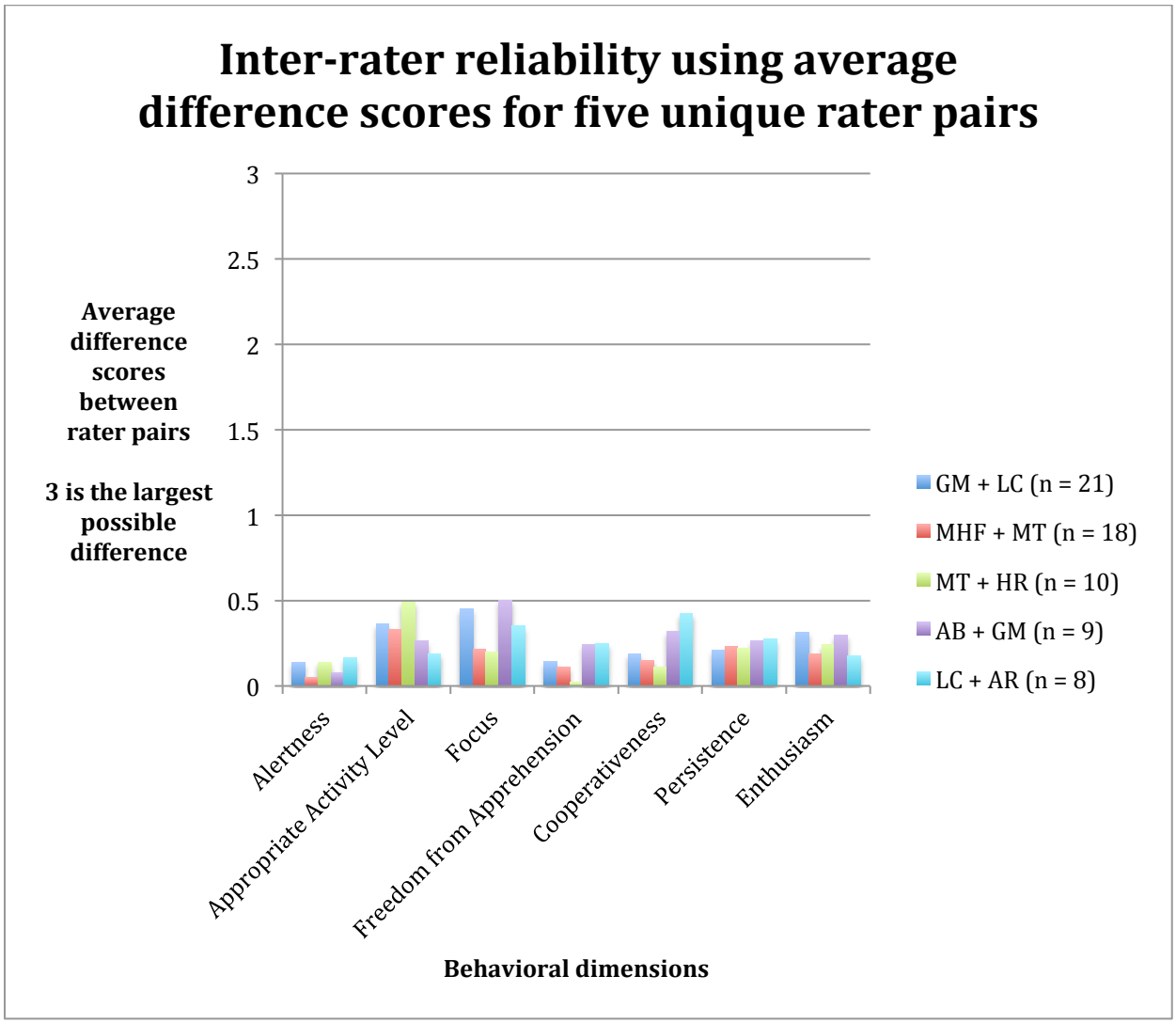
Wills, T. A., DuHamel, K., & Vaccaro, D. (1995). Activity and mood temperament as predictors of adolescent substance use: test of a self-regulation meditational model. *Journal of Personality and Social Psychology*, *68*, 901-916.

Windle, M. (1991). The difficult temperament in adolescence: associations with substance use, family support, and problem behaviors. *Journal of Clinical Psychology*, *47*, 310-315.

Appendix A

Graph 1

Inter-rater reliability assessment using average difference scores for five of the 16 unique rater pairs



Graph 2

Average behavior ratings examined by behavioral dimension for the female comparisons, females at genetic risk, male comparisons, and males at genetic risk

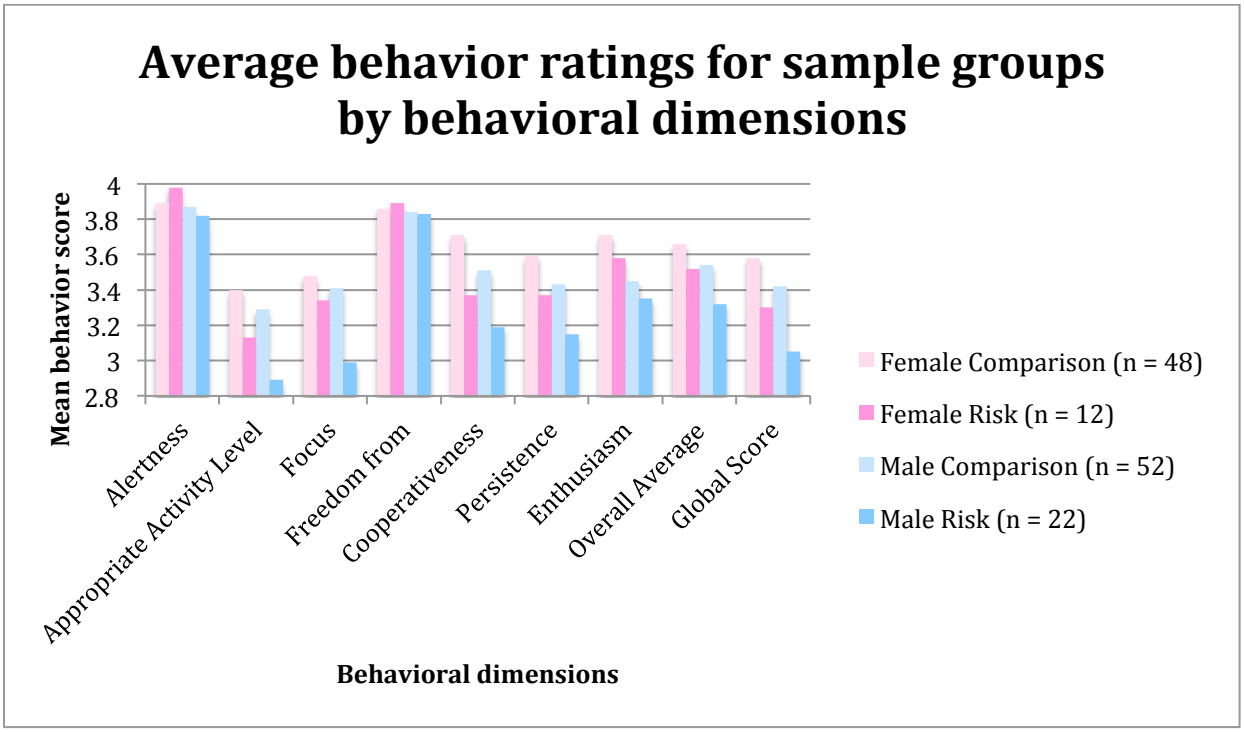


Table 1

Descriptions of the seven behavioral dimensions used to rate behavior

Behavioral dimension	**Description**
Alertness	Awake and alert—NOT sleepy or drowsy
Appropriate Activity Level	Appropriate for tasks, neither too low or too high
Focus	Attention directed to tester or materials as appropriate; not easily distracted but also not fixated on one activity or object
Freedom from Apprehension	Seems comfortable, does NOT appear anxious regarding test, environment, or test materials
Flexible Cooperativeness	Child responds to instructions of tester and/or parent; moves easily and flexibly from one activity to another as prompted
Persistence	Exhibits effort and continues trying, as necessary; not easily frustrated
Enthusiasm	Affect reflects interest in activity or materials; does not need extensive encouragement

Table 2

Comparison of average behavior ratings for comparison group (C) and at genetic risk group (R)

	C group (n = 100)		R group (n = 34)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.88	0.22	3.88	0.26	-0.11	*.915*	0
Activity Level	3.34	0.50	2.98	0.74	-2.68	***.010***	0.57
Focus	3.44	0.42	3.11	0.68	-2.65	***.011***	0.58
Freedom from Apprehension	3.85	0.27	3.85	0.27	0.05	*.961*	0
Flexible Cooperativeness	3.60	0.44	3.25	0.76	-2.55	***.015***	0.56
Persistence	3.51	0.41	3.23	0.70	-2.26	***.029***	0.49
Enthusiasm	3.57	0.47	3.43	0.58	-1.41	*.161*	0.27
Average rating of behavior	3.60	0.30	3.39	0.47	-2.73	*.019*	0.53

Table 3

Comparison of average behavior ratings for male group and female group

	Males (n = 74)		Females (n = 60)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.86	0.24	3.91	0.21	1.23	*.22*	0.22
Activity Level	3.17	0.65	3.34	0.49	1.75	*.083*	0.30
Focus	3.28	0.58	3.45	0.42	1.97	*.051*	0.33
Freedom from Apprehension	3.84	0.29	3.87	0.24	0.64	*.523*	0.11
Flexible Cooperativeness	3.41	0.62	3.64	0.44	2.46	***.015***	0.43
Persistence	3.35	0.55	3.55	0.44	2.27	***.025***	0.40
Enthusiasm	3.42	0.56	3.68	0.37	3.29	***.001***	0.55
Average rating of behavior	3.48	0.40	3.63	0.28	2.69	***.008***	0.44

Table 4

Comparison of average behavior ratings for male comparison group and female comparison group

	CM (n = 48)		CF (n = 52)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.87	0.21	3.89	0.23	0.38	*.703*	0.09
Activity Level	3.29	0.55	3.40	0.44	1.08	*.284*	0.22
Focus	3.41	0.48	3.48	0.35	0.88	*.380*	0.17
Freedom from Apprehension	3.84	0.29	3.87	0.24	0.38	*.704*	0.11
Flexible Cooperativeness	3.51	0.51	3.71	0.33	2.34	***.022***	0.47
Persistence	3.43	0.47	3.59	0.33	1.95	*.055*	0.39
Enthusiasm	3.45	0.54	3.71	0.34	2.97	***.004***	0.58
Average rating of behavior	3.54	0.33	3.66	0.24	2.08	***.040***	0.42

Table 5

Comparison of average behavior ratings for male at genetic risk group and female at genetic risk group

	RM (n = 22)		RF (n = 12)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.82	0.31	3.98	0.06	2.20	***.038***	0.72
Activity Level	3.89	0.78	3.13	0.64	0.90	*.376*	1.07
Focus	2.99	0.69	3.34	0.61	1.45	*.156*	0.54
Freedom from Apprehension	3.83	0.30	3.89	0.21	0.61	*.547*	0.23
Flexible Cooperativeness	3.19	0.80	3.37	0.71	0.66	*.512*	0.24
Persistence	3.15	0.69	3.37	0.71	0.90	*.375*	0.31
Enthusiasm	3.35	0.62	3.58	0.48	1.09	*.283*	0.41
Average rating of behavior	3.32	0.50	3.52	0.40	1.21	*.234*	0.44

Table 6

Comparison of average behavior ratings for female comparison group and female at genetic risk group

	FC (n = 48)		FR (n = 12)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.89	0.23	3.98	0.06	-2.28	***.026***	0.54
Activity Level	3.40	0.44	3.13	0.64	1.71	*.093*	0.49
Focus	3.48	0.35	3.34	0.61	0.77	*.458*	0.28
Freedom from Apprehension	3.86	0.24	3.89	0.21	-0.40	*.693*	0.13
Flexible Cooperativeness	3.71	0.33	3.37	0.71	1.60	*.136*	0.61
Persistence	3.59	0.33	3.37	0.71	1.05	*.314*	0.40
Enthusiasm	3.71	0.34	3.58	0.48	1.10	*.277*	0.31
Average rating of behavior	3.66	0.24	3.52	0.40	1.17	*.263*	0.42

Table 7

Comparison of average behavior ratings for the male comparison group and male at genetic risk group

	MC (n = 52)		MR (n = 22)				
Measure	*M*	*SD*	*M*	*SD*	*t*(df)	*p*	Cohen's *d* effect size
Alertness	3.87	0.21	3.82	0.31	0.82	*.414*	0.19
Activity Level	3.29	0.55	2.89	0.78	2.15	***.040***	0.59
Focus	3.41	0.48	2.99	0.69	2.98	***.004***	0.71
Freedom from Apprehension	3.84	0.29	3.83	0.30	0.11	*.910*	0.03
Flexible Cooperativeness	3.51	0.50	3.19	0.80	1.74	*.092*	0.48
Persistence	3.43	0.47	3.15	0.69	2.09	***.040***	0.47
Enthusiasm	3.45	0.54	3.35	0.62	0.65	*.520*	0.17
Average rating of behavior	3.54	0.33	3.31	0.50	1.92	*.065*	0.54

Appendix B

Forms

Rate date _______ Binet Rater initials_____ 2nd Half Rater Initials_____ 2nd Half rate date__________ Child ID___________ TEST DATE __________ Visit ___ Age___

Ratings of Child Behavior during Stanford Binet Assessments V6 & V7

Behavioral state may compromise the validity of the scores on behavioral assessments. This tool is designed to collect objective ratings based on observed behaviors, captured on recordings, of the assessment session, to help determine the possible validity of scores obtained.

Please note: a) A behavioral dimension may vary within a subscale—use what you consider most representative of either the predominant score, or the averaged score.
b) TWINS may have different parts of their assessments on different DVDs, so it makes sense to score twins back to back, going through each DVD completely. BE CAREFUL to score twins on the correct form.

1. Low or Absent : The behavioral dimension is **rarely observed, that is, the level is very low or absent**.
2. Marginal/Mixed: The behavioral dimension is **seen only occasionally**.
3. Adequate: The behavioral dimension is **seen frequently enough to be considered adequate but not necessarily optimal** in supporting the child's performance.
4. Clearly Demonstrated: The behavioral dimension is **clearly and frequently demonstrated and appears to be at good to optimal levels** in supporting the child's performance.

	Nonverbal Routing (NR)	Verbal Routing (VR)	Nonverbal Visual-Spatial Processing (VS)	Nonverbal WM (NVWM)	Verbal Fluid Reasoning (VFR)	Verbal WM (VWM)
Alertness: awake and alert—NOT sleepy or drowsy	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Appropriate activity level: appropriate for tasks, neither too low or too high	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Focus: attention directed to tester or materials as appropriate; not easily distracted but also not fixated on one activity or object	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Freedom from apprehension: seems comfortable, does NOT appear anxious regarding tester, environment, or test materials	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Flexible cooperativeness with adult: child responds to instructions of tester and/or parent; moves easily and flexibly from one activity to another as prompted	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Persistence: exhibits effort and continues trying, as necessary; not easily frustrated	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
Enthusiasm: affect reflects interest in activity or materials; does not need extensive encouragement	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
GLOBAL rating of OVERALL impression of child behavior	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

NOTES:

___ Please check if parent was in room during assessment and his/her level of involvement in the assessment.

Manual

V6 Child Behavior Scoring Manual

Instructions

You will be watching videos of 4-year-old assessments and rating the child's behavior on a scale of 1 to 4, based on several dimensions. You will then create a global score and write notes about why you scored the behavior as you did. Lastly, you will check the box at the bottom of the page to indicate if the parent was in the room and add a brief description of parent involvement in the assessment (i.e. if parent were involved in emotional/behavioral regulation). Please write "no" if parent was not in the room.

Range of Ratings

4: Clearly Demonstrated: The behavioral dimension is **clearly and frequently demonstrated and appears to be at good to optimal levels** in supporting the child's performance.

3: Adequate: The behavioral dimension is **seen frequently enough to be considered adequate but not necessarily optimal** in supporting the child's performance.

2: Marginal/Mixed: The behavioral dimension is **seen only occasionally**.

1: Low or Absent: The behavioral dimension is **rarely observed, that is, the level is very low or absent**.

Behavioral Dimensions

Alertness: How awake or drowsy the child appears to be. Signs of drowsiness are rubbing eyes, slouching in seat, resting body on table, yawning, and saying, "I'm tired." Typically, particularly with this behavioral dimension, a low alertness rating will impact other behavioral dimensions greatly (i.e. a child will rarely be exhausted and focused).

4. Child seems energetic and shows no signs of sleepiness
3. Child seems primarily energetic but shows a few signs of sleepiness that do not seem to compromise other behavioral dimensions
2. Child shows consistent signs of sleepiness throughout section that seem to compromise other behavioral dimensions with only a few moments of liveliness
1. Child rarely seems energetic, and sleepiness obviously compromises other behavioral dimensions

Appropriate activity level: Takes into account how physically active child is during assessment. Activity levels that are high *or* low can both create a testing environment that compromises optimal performance.

4. Child stays seated in chair with minimal fidgeting
3. Child fidgets/squirms or taps feet; may get out of chair once or twice
2. Child struggles to sit still (fidgets/squirms excessively and/or gets out of chair more than once or twice)
1. Child is unable to sit still; may spend more time wandering around room than in seat

Focus: How well attention is directed to tester or materials as appropriate (not playing with materials, but using them for their intended purpose); not easily distracted but also not fixated on one activity, object, or person in room.

4. Child looks at or attends to tester or test material consistently
3. Child frequently attends to tester/test material, but attention wanders at moments

2. Child more often than not is distracted by other items in room but attends to tester/test materials at moments
1. Child is unable to appropriately attend to tester/test materials

Freedom from Apprehension: seems comfortable and not excessively shy; does NOT appear anxious regarding tester, environment, or test materials. Signs of apprehension include avoiding eye contact, speaking quietly, looking to parent for comfort frequently, covering mouth, seeming hesitant to answer.

4. Child does not seem shy or uncomfortable at all
3. Child seems slightly shy/uncomfortable, but freely answers anyway; child may start out uncomfortable but may warm up throughout the section
2. Child's discomfort interferes with ability to participate, but child gives some form of an answer
1. Child's discomfort almost completely prevents participation

Flexible **cooperativeness** with adult: child responds to instructions of tester and/or parent; moves easily and flexibly from one activity to another as prompted (please note that testers occasionally allow child to remain out of seat while testing, which should not be considered uncooperative UNLESS tester does so in response to child's inability to stay in seat).

4. Child consistently follows tester/parent instruction
2. Child resists tester/parent instruction but complies after 1 or 2 prompts
3. Child resists instruction and requires many prompts from tester/parent but DOES eventually comply
1. Child consistently does not (or refuses to) comply with tester/parent instruction

Persistence: exhibits effort and continues trying, as necessary; not easily frustrated. Persistence includes child's willingness to elaborate on verbal answers, lack of impulsivity when responding, and determination to give correct answers.

4. Child seems to be giving best effort willingly
3. Child seems to put forth effort but may need a bit of prompting
2. Child is reluctant to give answers or put forth effort; may give one-word answers, give up on visual-spatial tasks after one try, or put little thought into answers
1. Child puts almost no effort into responses; requires extensive prompting and still does not seem to try

Enthusiasm: affect reflects interest in or excitement about testing situation.

4. Child has positive affect and seems excited to participate
3. Child frequently seems excited to participate but shows negative affect occasionally
2. Child shows negative affect more frequently than positive affect
1. Child primarily has a negative affect

Global Scores

After you have rated each behavioral dimension individually, you will decide on a global score. The global score is an overall impression of a child's behavior. The ultimate question you want

to ask yourself is if the child created an environment for him/herself that was conducive to his/her optimal performance on these tasks.

- The global score heavily considers the context, while the individual scores are purely observational. Was this behavior appropriate for a testing environment? For example, if a child is wandering around the room, he/she would score low on activity level, but if a tester allowed the child to wander while testing, and this behavior did not affect performance, then the child would still score high globally.
- The global score is NOT necessarily an average. A child may score high in a number of dimensions but very low in others that hold more weight (i.e. a child may be incredibly enthusiastic, alert, and free from apprehension, but he/she may be off topic or distracted by other items in the room and unable to stay in seat or cooperate—this type of child would have a low global score).
- Consider how hard the child made the tester work to keep child on task and cooperating. If a tester is jumping through hoops, the child will likely score low globally.

Keep in Mind

- Four-year-olds are young and frequently are unaccustomed to sitting still and focusing for such a long period of time. These ratings should be relative to a typical four-year-old—not to how a college student would behave in an assessment.
- The behavioral dimensions are hard to tease apart. Frequently, one will affect the others. Try to judge each one separately and identify the one that may be causing the others to be lower.
- As the child gets further into each section, the test becomes harder. Try not to mistake inability to complete tasks for lack of cooperation.

Training Example

Nonverbal Routing	
Alertness	Child seems very alert. **Circle 4.**
Activity Level	Activity level is appropriate for setting. Child is slightly fidgety, but does not disrupt testing environment. **Circle 4.**
Focus	Child is paying attention and looking at test booklet. Not distracted by other items/people in the room. **Circle 4.**
Freedom from Apprehension	Child seems comfortable with tester—not shy. Testing situation does not cause child anxiety. **Circle 4.**
Flexible Cooperativeness	Child follows tester's instructions. Tester does not need to direct child to stay on task. **Circle 4.**
Persistence	Child seems to keep trying until he decides on an answer. **Circle 4.**
Enthusiasm	Child shows interest in and interacts with testing materials (ex. Says, "Meow!"). **Circle 4.**
Nonverbal Routing Global	Child's behavior is very appropriate for testing environment. Stays focused and engaged. **Circle 4.**
Verbal Routing	
Alertness	Child seems alert. Rubs eyes, but seems to be an issue of boredom, rather than alertness. **Circle 4.**
Activity Level	Child is fairly fidgety—puts legs in shirt and flops around. Slightly inappropriate activity level for this situation, but does not seem to get in the way of participation. Activity level could be an effect of enthusiasm, rather than the underlying problem with behavior. **Circle 3.**
Focus	Child stays in his seat and participates. Attention is directed towards tester and materials, rather than other objects/people in the room. **Circle 4.**
Freedom from Apprehension	Child does not hesitate to express himself to the tester. Does not seem shy at all. **Circle 4.**
Flexible Cooperativeness	Child is slightly noncompliant, but tester is able to convince him to participate. **Circle 3.**
Persistence	Child does not elaborate when tester prompts him for more information. Even though may not be trying his hardest, he still attempts to answer questions. **Circle 3.**
Enthusiasm	Child clearly expresses disinterest throughout the duration of activity (ex. Says, "I don't want to do this."). Not crying, but not happy to participate. **Circle 2.**

Verbal Routing Global	Even though child was unenthusiastic and lacked slightly in persistence, he participated, and his behavior was conducive to testing. **Circle 3.**
Nonverbal Visual-Spatial Processing	
Alertness	Child seems awake and alert. **Circle 4.**
Activity Level	Child is slightly fidgety, but does not seem to affect ability to perform. Some of his movement even seems to stem from excitement/engagement. **Circle 4.**
Focus	Child stays in his seat and participates. Attention is directed towards tester and materials, rather than other objects/people in the room. **Circle 4.**
Freedom from Apprehension	Child does not hesitate to express himself to the tester. Does not seem shy at all. **Circle 4.**
Flexible Cooperativeness	At times, child is very cooperative. However, a few times, child expresses disinterest, but after tester and mother encourage him several times, he participates. On average, his cooperativeness could be better, but does not get too in the way of the test. **Circle 3.**
Persistence	Child attempts to build shapes until they are complete. In between shape people, child resisted participation, but once he started construction, he put forth effort until they were complete. **Circle 4.**
Enthusiasm	Enthusiasm fluctuated greatly. Child was very enthusiastic at times (ex. Says, "This is my favorite!"). At other times, he seems to dislike the task (ex. Says, "Just one more.") and needs extensive encouragement. Because this dimension is so debatable, we chose to score lower. **Circle 2.**
Nonverbal Visual-Spatial Global	Child resisted participation and was less enthusiastic as the test went on, but tester and mother convinced him to complete the task. Behavior was neither completely conducive nor detrimental to testing environment. **Circle 3.**
Nonverbal Working Memory	
Alertness	Child seems awake and alert. **Circle 4.**
Activity Level	Activity level was appropriate for testing situation. **Circle 4.**
Focus	Child lost focus a couple times (distracted by buried treasure and the back of his chair), but tester was able to

	redirect his focus. **Circle 3.**
Freedom from Apprehension	Child does not hesitate to express himself to the tester. Does not seem shy at all. **Circle 4.**
Flexible Cooperativeness	Child resisted very slightly, but did not require encouragement. **Circle 4.**
Persistence	Child exhibits sufficient effort throughout test. **Circle 4.**
Enthusiasm	Child seemed neutral about game. Did not express interest or disinterest very much. **Circle 4.**
Nonverbal Working Memory Global	Though child's focus strays only slightly, overall his behavior in this section was appropriate for situation. **Circle 4.**
Verbal Fluid Reasoning	
Alertness	Child seems awake and alert until the last activity, at which point he seems to become slightly drowsy. May impact performance a bit. **Circle 3.**
Activity Level	Child become fairly fidgety right before the last activity, but during each task seemed to have an appropriate level of activity. **Circle 3.**
Focus	Focus was fine during the first two activities, but wanes right before and during the silly sentences (ex. Child is preoccupied by mother's foot and seems to zone out a bit). Focus is not perfect but is reasonably acceptable given the task and his age. **Circle 3.**
Freedom from Apprehension	Child is not apprehensive or shy at all. **Circle 4.**
Flexible Cooperativeness	Tester has to convince child in between activities and a little bit during activities to participate, but child's noncompliance does not get in the way of him completing activities or put excessive strain on the tester. **Circle 3.**
Persistence	Persistence was fine during the first two activities. During silly sentences, persistence falters a bit (likely due to the difficulty of this level). Overall, persistence was acceptable for his age and the test items. **Circle 3.**
Enthusiasm	Enthusiasm was low for the duration of VFR. Although he completed items, he expressed disinterest throughout section. **Circle 2.**
Verbal Fluid Reasoning Global	Child was unenthusiastic and slightly noncompliant but completed activities. His behavior could have been better and may have negatively impacted tester's ability to assess him to a small extent, but it was not detrimental to the testing environment. **Circle 3.**

Verbal Working Memory	
Alertness	Child seems awake and not drowsy. **Circle 4.**
Activity Level	Child was rocking in chair and very fidgety, which was disruptive to the test. Not appropriate activity level for a testing situation. **Circle 2.**
Focus	While child was participating in repeating sentences, his focus was strong. **Circle 4.**
Freedom from Apprehension	Child is not apprehensive or shy at all. **Circle 4.**
Flexible Cooperativeness	Child was resistant to repeating sentences but did complete the task. However, he was so noncompliant before the next task that tester had to end the test before the second activity. **Circle 1.**
Persistence	Child put forth effort when he was repeating sentences. His noncompliance was more the cause of him not completing the section, rather than a lack of persistence. **Circle 4.**
Enthusiasm	Child complained about completing both parts of this section, and he complained so much about the second task that the tester was unable to convince him to do it (ex. Child screeched when asked to participate). **Circle 1.**
Verbal Working Memory Global	Child was uninterested in section and so uncooperative that he did not complete the test. Behavior was extremely inappropriate for testing situation. **Circle 1.**