POSTNATAL TESTOSTERONE AND AUTISTIC TRAITS IN 4- TO 7-YEAR-OLD CHILDREN

A Dissertation

by

JANET SAENZ

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Gerianne M. Alexander
Committee Members,	Robert W. Heffer
	Teresa Wilcox
	Jennifer B. Ganz
Head of Department,	Douglas Woods

December 2014

Major Subject: Psychology

ABSTRACT

Previous research has established a relationship between prenatal testosterone and sex-linked behavior, such that higher levels of prenatal testosterone are related to more masculine behaviors. Yet, little is known about the postnatal androgen surge, that occurs around three-months of age. Preliminary research suggests that higher levels of postnatal testosterone may also result in the masculinization of some sex-typed behaviors (e.g., visual preferences and toy preferences). The primary purpose of this study was to gain a better understanding of the relationship between postnatal testosterone and behaviors in early childhood, including behaviors characteristic of Autism Spectrum Disorders.

Sixty-six children, between 4- and 7-years of age, and their parents participated in this study. Saliva samples were obtained and assayed for testosterone levels when children were 3- to 4-months of age. Second to fourth digit ratios were measured and used as indicators of prenatal testosterone exposure. Parents completed one self-report questionnaire (Autism Spectrum Quotient) and several parent-report questionnaires (Autism Spectrum Quotient-Children's Version, Empathy Quotient-Children's Version, Systemizing Quotient-Children's Version, and Child Behavior Checklist), while children participated in various experimental tasks (Expressive One Word Picture Vocabulary Test, Playmate and Play Style Preferences Structured Interview, mental rotation task, two false belief tasks, an emotion recognition task, and an eye-tracking task).

Overall, lower digit ratios, indicative of greater prenatal testosterone exposure, were associated with more masculine behaviors. For example, lower digit ratios were

predictive of less time required to complete a mental rotation task and were correlated with lower empathy scores, smaller vocabularies, lower emotion recognition scores and less visual interest in viewing the eye portion of faces. A novel finding of this study was that higher postnatal testosterone levels at 3-months of age were predictive of less visual interest in people compared to objects at 5-years of age. This finding supports the idea that the postnatal surge of testosterone that occurs in early infancy could have organizational effects on behavior, especially behaviors required for socialization. However, future research establishing a link between postnatal testosterone and behaviors beyond early childhood is required before establishing this postnatal phase as another critical period for the development of later behaviors.

DEDICATION

Dedicated to my son and my husband. I hope that our family can continue to foster a passion for learning and an appreciation for higher education. With much love!

-Janet

ACKNOWLEDGEMENTS

First, I would like to thank my committee chair, Gerianne Alexander, for her guidance and support. I will eternally be grateful to her for accepting me into her lab and believing in me even when I doubted myself. I admire her expertise in the area of hormones and behavior, as well as her leadership and professionalism. Thank you Gerianne for being an excellent mentor!

I would also like to thank my committee members for the energy they have invested in this dissertation and their suggestions to improve the project. In addition, I would like to thank my previous instructors, supervisors, and colleagues at Texas A&M for making my journey through graduate school an enjoyable process. Similarly, I would like to thank my training director and supervisors at Children's Hospital Los Angeles for a wonderful internship experience.

Last but not least, I would like to thank my family for their encouragement and inspiration. I would especially like to thank my husband for all the hours we spent bouncing ideas off of one another and the countless cups of coffee that we shared. I am also appreciative of our family's help with parenting our little one.

NOMENCLATURE

ASD	Autism Spectrum Disorder
EMB	Extreme Male Brain
2D:4D	Second and Fourth Digits

TABLE OF CONTENTS

Page

LIST OF TABLES

1. INTRODUCTION

1.1 Testosterone

Sex hormones, such as testosterone, are important for the sexual differentiation of male and female phenotypes and behavior (Arnold, 1996; Breedlove, 1994; Hines, 2004). Experimental studies have demonstrated that animals exposed to higher levels of testosterone during periods of perinatal development exhibit more masculine behaviors (Arnold & Gorski, 1984; Phoenix, Goy, Gerral, & Young, 1959), while those exposed to lower levels of testosterone during periods of perinatal development exhibit less masculine behaviors (MacLusky & Naftolin, 1981). Although experimental studies of prenatal hormone effects have not been conducted in humans for obvious ethical reasons, support for hormone-dependent masculinization and defeminization of behavior comes from studies of individuals with increased androgens during utero because of endocrine disorders (Hines & Kaufman, 1994; Pasterski et al., 2011).

Sex hormones fluctuate through the lifespan and in males increase at three timepoints in development (once prenatally and twice postnatally). The prenatal surge of testosterone in males occurs from week 8 through week 24 of gestation (Breedlove, Cooke, & Jordan, 1999; Hines, 2002). Testosterone levels during this early developmental phase are thought to have organizational effects on brain development, such that they produce permanent changes in brain structures and function that shape later personality and behavior (Peper & Koolschijn, 2012). The second surge of testosterone in males occurs from 1 month through 3 months of age (Forest, Sizonenko,

Cathiard, & Bertrand, 1974). During this transient activation of the hypothalamic-pituitary-axis, testosterone levels rise to adult levels before falling to low childhood levels where they remain until puberty, when the third surge of testosterone occurs (Baron-Cohen, Lutchmaya, & Knickmeyer, 2004). Testosterone levels in adolescents and adults are thought to produce activational or transient effects on behavior that are dependent on the presence of critical levels of hormone and on prior organizational effects (Buchanan, Eccles, & Becker, 1992).

Prenatal testosterone levels can be measured in amniotic fluid and umbilical cord blood. Of these two methods the expected sex differences in androgen levels are documented more consistently in amniotic fluid samples. However, amniotic fluid samples are often risky to obtain. A proxy measure of prenatal testosterone exposure that is easier to obtain and non-invasive is the ratio of the length of the second digit (index finger) over the length of the fourth digit (ring finger) (2D:4D) (J. T. Manning, Bundred, Newton, & Flanagan, 2003). Larger 2D:4D ratios are thought to be indicative of less prenatal androgen exposure, while smaller 2D:4D ratios are thought to be indicative of greater prenatal androgen exposure (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004). Despite some controversy surrounding this method (Berenbaum, Bryk, Nowak, Quigley, & Moffat, 2009), the inverse association between 2D:4D ratios and prenatal androgen exposure has been supported by findings indicating a negative relationship between fetal testosterone levels and 2D:4D ratios at the age of two-years (Lutchmaya et al., 2004), as well as studies indicating that 2D:4D ratios are smaller (more masculine) in females exposed to abnormally high levels of prenatal androgens

due to endocrine disorders (W. M. Brown, Hines, Fane, & Breedlove, 2002; Okten, Kalyoncu, & Yaris, 2002). Although 2D:4D ratios can be measured in both hands, right hand 2D:4D ratios are often favored because they produce greater sex differences (J. Hönekopp & Watson, 2010).

Another approach used to investigate the effects of prenatal exposure to high testosterone levels is to compare healthy individuals to individuals born with endocrine disorders, such as congenital adrenal hyperplasia (CAH). CAH is a genetic disorder that results in an inability to produce the adrenal hormone cortisol and overproduction of androgens, such as testosterone (Hines, 2004; Pang et al., 1980). Females with CAH exhibit more masculine patterns of behavior than their unaffected siblings. For example, girls with CAH are more likely to select boy playmates (Hines & Kaufman, 1994), more likely to prefer playing with male-typical toys (Berenbaum & Hines, 1992; Pasterski et al., 2005), and more likely to be rated as aggressive (Berenbaum & Resnick, 1997; Pasterski et al., 2007).

Postnatal testosterone is measured by extracting circulating testosterone levels from saliva and serum samples (Baron-Cohen et al., 2004). For practical reasons, saliva samples are most often used in research with young children. Higher levels of postnatal testosterone have been associated with the masculinization of sex-typed behaviors. For example, higher levels of postnatal testosterone in children have been associated with more masculine visual preferences, emotion regulation, and play styles (Alexander & Saenz, 2011; Alexander, Wilcox, & Farmer, 2009; Lamminmäki et al., 2012).

In sum, sex hormones are important for the sexual differentiation of psychological characteristics. Sex hormones exert organizational and activational effects, they reach peak levels in males during 3 times in development, and can be measured in a number of ways. It is also important to highlight that although the effects of atypical hormone levels can be measured by contrasting behavior in relevant clinical samples and healthy individuals, it is more efficient to study the normal variability in hormone levels within healthy individuals.

1.2 Testosterone and psychopathology

Sex differences exist in the natural history and prevalence rates of many psychological disorders. Boys are more likely than girls to experience psychological disorders in early childhood (Meltzer, Gatward, Goodman, & Ford, 2003). Conversely, girls are more likely than boys to experience psychological disorders in adolescence and adulthood (Hartung & Widiger, 1998). Typically prepubertal boys are more often diagnosed with externalizing disorders (i.e., conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder) and postpubertal girls are more often diagnosed with internalizing disorders (i.e., depression, anxiety, and eating disorders) (Zahn-Waxler, 1993). These differences in the prevalence rates and trajectories of mental disorders in boys and girls have been explained by various theories, including socialization, genomic imprinting, sex-linked brain development, and sex-differential selective pressures (Baron-Cohen et al., 2004; Crespi & Badcock, 2008; Zahn-Waxler, 1993).

One overarching biological theory that attempts to explain psychological disorders that show a male prevalence is the Extreme Male Brain (EMB) theory (Baron-Cohen, 2002, 2010). This theory is an extension of the Empathizing-Systemizing (E-S) theory (Baron-Cohen, 2002) which states that empathizing and systemizing traits are present in both sexes, but females are innately superior to males at identifying and responding to other's thoughts and emotions (empathizing) and males are innately superior to females at constructing systems and identifying variables and rules within a system (systemizing). Thus, females are typically better able to predict a person's behavior and males are better able to predict a system's behavior. Similarly, the EMB theory proposes that higher levels of fetal testosterone organize brain structures that support systemizing and other male-typical behaviors that are associated with male prevalent disorders. In other words, fetal testosterone is believed to masculinize some behaviors and personality traits. Although this theory has been primarily applied to explain autism as an extreme manifestation of sexually dimorphic characteristics it has also been applied in explanations of other psychological disorders that are male dominant.

Although the original EMB theory did not propose an "extreme female brain" as a determinant of psychological disorders more prevalent in females, this idea has been more recently proposed. Specifically, it has been suggested that having an extreme female brain could also be problematic and related to psychopathology, since having more female qualities (e.g., higher levels of empathy) has been linked to internalizing disorders during adolescence (Zahn-Waxler, Shirtcliff, & Marceau, 2008). Another

hypothesis to explain why disorders more commonly diagnosed in females (e.g., mood disorders, anxiety disorders, eating disorders) arise during puberty asserts that organizational effects that transpired early in development are activated during puberty, resulting in sex differences in vulnerability to these psychological disorders (Peper & Koolschijn, 2012).

Some support for the idea that higher fetal testosterone levels result in a greater risk for male-typical disorders, while lower fetal testosterone levels results in a greater risk for female disorders comes from studies using 2D:4D ratios. Smaller 2D:4D ratios, which are indicative of greater prenatal testosterone exposure, have been related to less prosocial behaviors and greater probability of having schizophrenia in females (Fink, Manning, Williams, & Podmore-Nappin, 2007; Venkatasubramanian, Arasappa, Rao, & Gangadhar, 2011), more conduct problems in males (Fink et al., 2007), and more autistic traits and hyperactivity among males and females (Fink et al., 2007; Milne et al., 2006; Williams, Greenhalgh, & Manning, 2003). Larger 2D:4D ratios, which are indicative of less prenatal testosterone exposure, have been related to emotional symptoms and anxiety in males (Evardone & Alexander, 2009; Williams et al., 2003) and bulimia in females (Quinton, Smith, & Joiner, 2011).

High levels of prenatal testosterone have also been implicated in the development of disorders of impulse control (i.e., tourette syndrome, obsession compulsive disorder) which show greater prevalence rates in boys than girls (Alexander & Peterson, 2001). Females with these tic-related disorders scored better on spatial tasks that are traditionally categorized as masculine than their unaffected relatives and both males and

females with tic-related disorders indicated greater preferences for male-typical play than their unaffected relatives (Alexander & Peterson, 2004). Interestingly, males with tic disorders who have presumably been exposed to high levels of prenatal testosterone scored worse on measures of spatial ability that are traditionally categorized as masculine than their unaffected relatives (Alexander & Peterson, 2004). However, this is not surprising since research indicates a curvilinear relationship between testosterone and spatial ability, such that while higher levels of prenatal testosterone in females strengthen spatial skills and other masculine cognitive abilities, very high levels of prenatal testosterone in males impair performance on these tasks (Berenbaum, Bryk, & Beltz, 2012; Vuoksimaa et al., 2010).

In conclusion, there is preliminary evidence linking testosterone levels to psychopathology. In particular, high levels of prenatal testosterone have been linked to disorders more prevalent in males, while lower levels of prenatal testosterone have been linked to disorders more prevalent in females. However, it is important to acknowledge that this theory of prenatal testosterone is not comprehensive and there likely are other factors (i.e., genes and environmental factors) that also contribute to the risk of psychopathology (Baron-Cohen et al., 2011).

1.3 Prenatal testosterone and autistic traits

According to the Diagnostic and Statistical Manual of Mental Disorders IV-TR (DSM-IV-TR), autistic disorder (autism) is a pervasive developmental disorder characterized by a "qualitative impairment in social interactions", "quantitative

impairment in communication", and "restricted repetitive and stereotyped patterns of behavior, interests, and activities" (American Psychiatric Association, 2000). Additionally, a delay in either social interactions, language, or imaginative play must be present prior to three years of age (American Psychiatric Association, 2000). The criteria for high functioning autism (Asperger's disorder) are the same except that there is no delay in language or communication necessary. Autism and Asperger's are diagnosed most often in males, with ratios of 4:1 and 11:1 respectively (Chakrabarti & Fombonne, 2005; Gillberg, Cederlund, Lamberg, & Zeijlon, 2006). These disorders have a strong genetic component and are often seen among fathers and brothers of children with an autism spectrum disorder (Eriksson, Westerlund, Anderlid, Gillberg, & Fernell, 2012; Hoekstra, Bartels, Verweij, & Boomsma, 2007). Although early interventions are usually helpful in targeting and ameliorating deficits observed in autism, there is currently no cure.

While the DSM-IV-TR was based on a model that classified disorders on a categorical basis, there has been growing recognition that the expression of autism characteristics lie on a continuum from very severe in individuals with autistic disorder to less impairing symptoms displayed in Asperger's disorder. For this reason these two disorders are collectively referred to in the research literature as Autism Spectrum Disorders (Karabekiroglu et al.). Recently, the new DSM (DSM-5) eliminated the distinctive diagnoses of autistic disorder and Asperger's disorder and replaced them with a new diagnosis (autism spectrum disorder) that has three levels of severity and is intended to parallel increasing evidence that psychopathology may be an extreme

presentation of normative traits (American Psychiatric Association, 2013). Since autism characteristics fall on a continuum from very present to absent, we can expect these characteristics to be present at varying degrees within typically developing individuals.

When investigating autism spectrum disorders and other psychological disorders researchers often take advantage of the fact that similar, albeit less severe, manifestations of the behaviors/symptoms expressed in those psychological disorders are present within healthy individuals. This allows them to study the normal variability of these behaviors within the general population and make inferences of their findings to the ASD community (e.g., analogue studies). This method of investigating psychological disorders is often utilized since it removes the difficulties involved in recruiting a large sample size from disorders that have low prevalence rates. Researchers applying this method of investigation to inform their understanding of ASD could study variability in the expression of social communication and repetitive-stereotyped behaviors within healthy children. A brief review of research on the interaction between prenatal testosterone levels (measured by direct and indirect methods) and the expression of autistic traits (in healthy individuals, individuals with an ASD, and individuals with CAH) is provided in the subsequent paragraphs.

As stated formerly, prenatal testosterone is believed to have masculinizing effects on sex-typed behaviors and to be implicated in the development of psychopathology. In particular, it has been proposed that ASD is an extreme manifestation of the sexually dimorphic traits outlined in the ASD diagnostic criteria (Baron-Cohen, 2002). In

essence, higher levels of prenatal testosterone is viewed as a risk factor for ASD (Auyeung & Baron-Cohen, 2012).

Parents report that typically developing boys display more autistic traits than typically developing girls and that children with an ASD display more autistic traits than typically developing children on parent-report questionnaires (Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008). Additionally, parents report that girls with CAH, who have been exposed to higher levels of prenatal testosterone, display more autistic traits than their unaffected relatives (Knickmeyer et al., 2006). Furthermore, autism scores of healthy children, obtained by parental-report, have been positively associated to fetal testosterone levels measured in amniotic fluid (Auyeung, Baron-Cohen, et al., 2009).

An essential characteristic of ASD is low frequency of eye contact (Kirchner, Hatri, Heekeren, & Dziobek, 2011). In healthy participants, girls engage in more eye contact than boys since early in infancy (Leeb & Rejskind, 2004). Eye contact has also been found to have an inverse relationship with fetal testosterone levels measured in amniotic fluid, such that higher levels of fetal testosterone are related to less eye contact (Lutchmaya, Baron-Cohen, & Raggatt, 2002a). Boys and girls also have differential looking preferences for social and nonsocial stimuli. Female infants that are 1-day-old display a looking preference towards social stimuli (face), while boys display a looking preference towards nonsocial stimuli (mobile) (Connellan, Baron-Cohen, Wheelwright, Batki, & Ahluwalia, 2000). Among 1-year-old children, girls prefer to look at social stimuli (people conversing) and boys prefer to look at nonsocial stimuli (moving cars) (Lutchmaya & Baron-Cohen, 2002). Similarly, typically developing toddlers prefer to

look at social stimuli (kids doing yoga), while toddlers with an ASD prefer to look at nonsocial stimuli (geometric shapes) (Pierce, Conant, Hazin, Stoner, & Desmond, 2011).

As mentioned earlier, females are superior to males at identifying other's thoughts and emotions (empathizing) and males are superior to females at constructing and exploring systems (systemizing) (Auyeung, Wheelwright, et al., 2009; Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003). Furthermore, children and adults with an ASD obtain higher systemizing scores and lower empathizing scores when compared to healthy individuals (Auyeung, Wheelwright, et al., 2009; Baron-Cohen et al., 2003). Empathizing scores have also been found to have a negative relationship with fetal testosterone levels measured in amniotic fluid, such that higher levels of fetal testosterone are associated with lower empathizing scores among healthy children (Chapman, Baron-Cohen, Auyeung, Taylor, & Hackett, 2006). In contrast systemizing scores have a positive relationship with fetal testosterone levels measured in amniotic fluid, such that higher levels of fetal testosterone are associated with higher systemizing scores among healthy children (Auyeung et al., 2006).

Generally, women perform better on emotion recognition tasks than men (Golan, Baron-Cohen, & Hill, 2006). Adolescent boys with Asperger's syndrome perform worse on emotion recognition tasks than healthy adolescent boys (Peter Krajmer, Špajdel, Celec, & Ostatníková, 2011). Similarly, children with an ASD perform worse on emotion recognition tasks than healthy children (Golan, Baron-Cohen, & Golan, 2008). False belief tasks measure an individual's theory of mind (TOM), or ability to make inferences about other's thoughts, feelings, and intentions. Consistent with the general

findings that girls are better at recognizing and predicting others' thoughts and behaviors than boys, girls outperform boys in false belief measures (Knickmeyer & Baron-Cohen, 2006). Theory of mind is impaired in individuals with an ASD compared to typically developing children (Baron-Cohen, Leslie, & Frith, 1985).

From an early age girls have larger vocabularies than boys (Butterworth & Morissette, 1996; Fenson et al., 1994). Vocabulary size has also been found to have a negative relationship with fetal testosterone levels measured in amniotic fluid, such that higher levels of fetal testosterone are predictive of a smaller vocabulary size among healthy children (Lutchmaya, Baron-Cohen, & Raggatt, 2002b). Additionally, girls with CAH, who have been exposed to higher levels of prenatal testosterone, perform lower than unaffected girls on measures of oral language and verbal fluency (Inozemtseva, Matute, & Juárez, 2008). Typically children with an ASD experience delays in language acquisition and display less communicative behaviors with their parents compared to typically developing children (O'Neill & Happe, 2000).

One of the most robust sex differences in cognitive abilities is found in mental rotation tasks. Mental rotation tasks measure a person's ability to hold a mental representation of an object so that one can successfully predict what that object would look like if it were rotated. These objects can be in the form of two-dimensional or three-dimensional drawings. Males are generally better than females at solving mental rotation tasks during infancy (Moore & Johnson, 2008; Quinn & Liben, 2008), childhood (Hahn, Jansen, & Heil, 2010; Hoyek, Collet, Fargier, & Guillot, 2012), and adulthood (Voyer & Hou, 2006). Presumably males outperform females on these tasks in part because of

higher levels of prenatal testosterone. Further evidence for the relationship between higher prenatal testosterone exposure and mental rotation task performance comes from studies which have found that females with a male fraternal twin outperform women with a female fraternal twin (Heil, Kavsek, Rolke, Beste, & Jansen, 2011; Vuoksimaa et al., 2010). Additionally, girls with CAH, who have been exposed to higher levels of prenatal testosterone, outperform their unaffected sisters on mental rotation tasks (Berenbaum et al., 2012). Lastly, rate of response on mental rotation items has been positively associated with fetal testosterone levels measured from amniotic fluid in girls (Grimshaw, Sitarenios, & Finegan, 1995). If ASD is an extreme presentation of the male brain, then children with an ASD should show extreme presentations of other male-typical behaviors, such as spatial ability and play styles.

A behavior that often shows a large sex difference in typically developing children is play style. Typically developing boys of all ages prefer to play with peers who are engaged in masculine play styles, regardless of their sex, while younger girls prefer to play with other girls, regardless of the play style (masculine or feminine) (Alexander & Hines, 1994). Research indicates that girls with CAH, who are exposed to higher levels of prenatal testosterone, express more masculine playmate preferences and more masculine play styles than their unaffected relatives (Pasterski et al., 2011). Play is often impaired in children with an ASD (Lewis, 2003). However, play style and playmate preferences have not been investigated in these children.

Overall, this review supports the masculinizing effect that prenatal testosterone exerts on autistic behaviors. It is also important to keep in mind that many of the

behaviors discussed in the former paragraphs are not independent from one another. For example, an early interest in biological motion, social stimuli, and eye contact are believed to be precursors to recognizing emotional states in others, developing a theory of mind, and displaying empathizing skills (Bal et al., 2010; Klin, Lin, Gorrindo, Ramsay, & Jones, 2009). Additionally, language development has been positively associated with the development of play and theory of mind (Farhadian, Abdullah, Mansor, Redzuan, & Kumar, 2010; Lewis, 2003; Lind & Bowler, 2009). Furthermore, the ability to detect systems and understand the rules of a system (systemizing) is predictive of performance on mental rotation tasks for men (C. M. Cook & Saucier, 2010).

1.4 Postnatal testosterone and behavior

In contrast to the relationship between prenatal testosterone and behavior, little is known about the relationship between the postnatal surge of testosterone, which occurs around 3- months of age, and behavior. Animal studies indicate that male rats that are prevented from undergoing the postnatal testosterone surge, by undergoing castration immediately after birth, do not experience a demasculinization of the corpus callosum, implying that the postnatal surge is not critical for the masculinization of the corpus callosum (Mack, McGivern, Hyde, & Denenberg, 1996). However, the postnatal surge of testosterone has been associated with the circadian rhythm in male rats (Zuloaga, McGivern, & Handa, 2009). Postnatal testosterone in rats is also important for the size of male reproductive organs (Macleod et al., 2010; van den Driesche et al., 2011). In non-

human primates, postnatal testosterone levels have been shown to be important for the normal development of male and female external genitalia (G. R. Brown, Nevison, Fraser, & Dixson, 1999). Additionally, suppression of postnatal testosterone levels in male primates has been shown to increase proximity seeking to their mother, a more female-typical behavior (Wallen, Maestripieri, & Mann, 1995). However, postnatal testosterone levels in male and female primates do not appear to affect play or mounting behavior (G. R. Brown & Dixson, 1999; Nevison, Brown, & Dixson, 1997).

In humans as in other primates, postnatal testosterone levels in infancy have been implicated in the normal development of male phallic and scrotal development (Main, Schmidt, & Skakkebæk, 2000). However, in contrast to findings in other mammals that suggest that the postnatal surge in males does not influence the sexual differentiation of behavior, recent human studies indicate that postnatal testosterone levels in 4-month-old infant boys are predictive of looking preference for male typical stimuli (a group of figures instead of a solitary figure) (Alexander, Wilcox, & Farmer, 2009). Postnatal testosterone levels in 4-month-old infant boys are also predictive of greater negative affect (Alexander & Saenz, 2011). Furthermore, postnatal testosterone levels have also been associated with lateralization patterns during an auditory phoneme processing task among 4-week-old males (Friederici et al., 2008). Moreover, postnatal testosterone scores obtained by averaging urinary samples collected once a month for the first six months of life have been positively correlated to gender role behavior in 14-month-old boys, positively correlated to the percentage of time girls played with a train, and

inversely correlated to the percentage of time boys played with a doll (Lamminmäki et al., 2012).

Preliminary evidence supports the hypothesis that early postnatal life may be another critical period for the development of gender-linked behavior. However, it remains unclear whether these findings reflect temporary and transient effects of postnatal hormones on behavior or if postnatal androgen will have long-term organizational effects, separate and independent from those of prenatal testosterone. If early postnatal life is indeed another critical period for the development of gender-linked behavior, then high postnatal testosterone levels may also be another risk factor for ASD and other disorders with sex-biased prevalence rates (e.g., conduct disorder, oppositional defiant disorder, attention deficit hyperactivity disorder).

1.5 Purpose of the current study

The purpose of this study was to increase our understanding of the relationship between the postnatal surge of testosterone and later patterns of behavior by examining whether postnatal testosterone levels collected at 3-months of age would be predictive of sex-typed behaviors in early childhood. It was hypothesized that after controlling for levels of prenatal androgen exposure (using 2D:4D ratios), greater postnatal testosterone levels in early infancy would be predictive of more male-typical behaviors in early childhood (i.e., higher autism quotients, lower empathy quotients, higher systemizing quotients, higher CBCL total problem scores, lower expressive language scores, higher mental rotation accuracy scores and lower reaction times, less time spent looking at

people and the eye region of faces, lower facial emotion recognition scores, lower theory of mind scores).

Although preliminary evidence of postnatal testosterone's masculinizing effect on behavior in early infancy does exists, no studies have examined the potential long-term effects that postnatal testosterone can have past the age of two years. Additionally, many of the behaviors investigated in this study are specific to autism spectrum disorders. Previous research has shown that the expression of ASD traits is influenced by greater prenatal testosterone exposure (Auyeung, Baron-Cohen, et al., 2009) and recent research has begun to investigate the effects of postnatal testosterone on ASD traits. To date, two studies have examined the effects of postnatal testosterone on ASD traits measured via questionnaire screeners at 18- to 24-months of age. The first study indicated that postnatal testosterone was not related to ASD traits; while the second study indicated that postnatal testosterone was predictive of ASD traits (Auyeung, Ahluwalia, et al., 2012; Saenz & Alexander, 2013b). This study is the first to conduct a comprehensive investigation of postnatal testosterone's influence on ASD traits.

2. METHODS

2.1 Participants

Participants included 66 children (35 male, 31 female) between 4- and 7-years of age (*M* = 1846 days, *SD* = 255 days; *M* = 5 years, 1 month, 16 days, *SD* = 8 months, 15 days), and their parents. These children were originally recruited through birth announcements and are part of a larger longitudinal study investigating the relationship between postnatal testosterone and sex-typed behavior. A previous study was conducted when these children were 3- to 4-months old and during that time saliva samples were obtained and assayed for testosterone levels. All former participants who were between 4- and 7-years of age were invited to participate in the current study. Signed, informed consent was obtained from all parents and verbal assent was obtained from all children.

2.2 Measures

2.2.1 Parent questionnaire

Autism Spectrum Quotient (AQ). The AQ is a 50-item, self-report questionnaire measuring the degree to which adults have traits associated with the autism spectrum (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). Five domains (social skill, attention switching, attention to detail, communication, and imagination) are measured with 10 items each. Respondents select 1 of 4 answer choices (“definitely agree”, “slightly agree”, “slightly disagree”, “definitely disagree”), which correspond to scores on a 4-point Likert scale (Austin, 2005; Hoekstra et al., 2007). This style of

scoring is used more frequently than the original binary method described in Baron-Cohen et al. (2001). This continuum-scoring style yields scores ranging from 0 to 150, with higher scores indicative of more autistic traits. The AQ can adequately differentiate between adults with Asperger syndrome or high-functioning autism and adults without these disorders (adults with Asperger's/autism score higher than typically developing adults), men and women without Asperger's/autism (men score higher than women), and students studying science, humanities, or social sciences (students studying science score higher than those studying humanities or social sciences). This instrument has moderate construct validity and good test-retest and interrater reliability.

2.2.2 Child questionnaires

Autism Spectrum Quotient-Children's Version (AQ-C). The AQ-C is a 50-item, parent-report questionnaire designed to measure autistic traits in children between the ages of 4- and 11-years (Auyeung et al., 2008). Similar to the adult and adolescent versions, this instrument assesses five areas (social skills, attention switching, attention to detail, communication, and imagination) with ten items per area. Respondents select one of four answer choices ("definitely agree", "slightly agree", "slightly disagree", "definitely disagree"), which correspond to scores on a 4-point Likert scale. Higher scores indicate more autistic traits. The AQ-C can adequately differentiate between children with Asperger syndrome or high-functioning autism and children without these disorders (children with Asperger's/autism score higher than children with no ASD) and boys and girls without Asperger's/autism (boys score higher than girls). This instrument

has excellent test-retest reliability, good construct validity, and satisfactory internal consistency for the five subscales.

Empathy Quotient-Children's Version (EQ-C). The EQ-C is a 27-item, parent-report questionnaire designed to measure the ability to recognize another person's emotions and thoughts in children between the ages of 4- and 11-years (Auyeung, Wheelwright, et al., 2009). The EQ-C was adapted from the Adult EQ (Baron-Cohen & Wheelwright, 2004). Respondents are asked to select 1 of 4 answer choices that correspond to scores on a 3-point Likert scale; 'definitely agree' corresponds to 2 points, 'slightly agree' corresponds to 1 point, while 'slightly disagree' and 'definitely disagree' correspond to 0 points. The point system is reversed for items that are worded in the reverse direction. Scores on the EQ-C range from 0 to 54, with higher scores indicative of greater empathy. The EQ-C can adequately differentiate between typically developing boys and girls (girls obtain higher scores than boys) and children with an ASD and typically children (typically developing boys obtain higher scores than children with an ASD). This instrument has good test-retest reliability and high internal consistency.

Systemizing Quotient-Children's Version (SQ-C). The SQ-C is a 28-item, parent-report questionnaire designed to measure the ability to analyze, explore, and construct a system in children between the ages of 4- and 11-years (Auyeung, Wheelwright, et al., 2009). The SQ-C was adapted from the Adult SQ (Baron-Cohen et al., 2003). Respondents are asked to select 1 of 4 answer choices that correspond to scores on a 3-point Likert scale; 'definitely agree' corresponds to 2 points, 'slightly agree' corresponds to 1 point, while 'slightly disagree' and 'definitely disagree'

correspond to 0 points. The point system is reversed for items that are worded in the reverse direction. Scores on the SQ-C range from 0 to 56, with higher scores indicative of greater systemizing abilities. The SQ-C can adequately differentiate between typically developing boys and girls (boys obtain higher scores than girls) and children with an ASD and typically developing children (children with an ASD obtain higher scores than typically developing boys). This instrument has good test-retest reliability and high internal consistency.

Child Behavior Checklist (CBCL/1.5-5). The CBCL/1.5-5 is a 99-item, parent-report questionnaire designed to measure behavioral and psychiatric problem in children between the ages of 1.5 and 5-years (Achenbach & Rescorla, 2000). Respondents are asked to select 1 of 3 answer choices that correspond to scores on a 3-point Likert scale; 'very true or often true' corresponds to 2 points, 'somewhat or sometimes true' corresponds to 1 point, while 'not true of the child' corresponds to 0 points. The CBCL/1.5-5 can be hand scored or computer scored, and provides 7 problem subscales (emotionally reactive, anxious/depressed, somatic complaints, withdrawn, attention problems, aggressive behavior, and sleep problems), two broad-band problem scales (internalizing problems and externalizing problems), and allows for the computation of a total problem score. Higher total problem scores indicate greater behavioral and psychiatric problems. The CBCL/1.5-5 has been used in numerous clinical and research settings; standard scores are used in clinical practice and raw scores in research settings. The CBCL/1.5-5 has acceptable internal validity, test-retest reliability, and predictive validity in American children.

2.2.3 Experimental tasks

Digit ratios. Photocopies of each child's right hand was obtained and the length of their ring and index fingers was measured from the basal crease to the tip of the finger by two independent raters using digital vernier calipers. The average ratio of the two ratings was used as a proxy measure of prenatal testosterone exposure. Higher ratios are indicative of less prenatal testosterone exposure and smaller ratios are indicative of greater prenatal testosterone exposure (J. T. Manning et al., 2003).

Sex-typed play styles. Playmate and Play Style Preferences Structured Interview (PPPSI) is a sex-typing task that asks children to identify their play style preferences and their playmate preferences (Alexander & Hines, 1994). During the interview children are presented with 40 cards, each depicting two figures playing with their own toy or activity, and asked to select which figure they would like to play with. Fourteen cards assess playmate preference by showing a girl and a boy playing with the same toy and asking children to select who they would prefer to play with. Thirteen cards assess play style preference by showing two gender-ambiguous figures playing with a masculine or feminine toy/activity and asking children to select who they would prefer to play with. Thirteen conflict cards depict a girl playing with masculine toys/activities and a boy playing with feminine toys/activities. This instrument provides a preference score for playmate and play style. Play style is made up of three subscales: toys (masculine vs. feminine), activity level (high vs. low), and play quality (rough-and-tumble vs. non-aggressive). This instrument shows a reliable sex difference in normative samples, with boys preferring to play with peers who engage in masculine play styles, regardless of

their sex, and younger girls preferring to play with other girls, regardless of the play style (masculine or feminine), and differentiates between girls exposed to higher levels of androgens and unaffected siblings (Alexander & Hines, 1994; Pasterski et al., 2005). Administration takes about 15 minutes.

Sex-linked Cognitive Tasks: (1) Vocabulary. Expressive One-Word Picture Vocabulary Test, Fourth Edition (EOWPVT-4) is a standardized test of expressive language designed for people between the ages of 2 and 80-years (Martin & Brownell, 2010). Administration consists of presenting colored pictures displaying an object, action or concept and asking the respondent to describe the picture in one word. Although this instrument does not show a reliable sex difference in normative samples, we hoped that it would relate to hormone levels (Stoner & Spencer, 1983). Administration typically takes about 7 to 15 minutes. Hand scoring takes approximately 5 minutes.

(2) Mental rotation ability. Children completed a simplified version of the original mental rotation task presented by Shepard and Metzler (1971). This task involved mental rotation of two-dimensional representations of panda bears (Marmor, 1975, 1977) and is used in previous research with children (Grimshaw et al., 1995; Perner, Kloo, & Rohwer, 2010). Participants underwent pre-training of the same/different concept and an experimental test. During pre-training children were taught to discriminate between same and different pairs of stimuli through explanation on 4 trials, each depicting two black and white, side-by-side bears standing in an upright position raising either their left arm (same), right arm (same), or one raising their right

arm and the other raising their left arm (different). The experimental test consisted of 20 trials consistently depicting the bear on the right side standing in an upright position while the bear on the left side rotated clockwise at a 0°, 45°, or 135° angle. A stopwatch was used to record each child's reaction time (from the time the stimuli are presented to the time they give an answer) and accuracy of answers was recorded. This task shows a reliable sex difference in normative samples, with boys generally outperforming girls (Grimshaw et al., 1995). Completion of this task takes approximately 10 to 15 minutes.

False Belief Tasks: (1) Location change task. False belief tasks were used to measure an individual's theory of mind (TOM) or one's ability to engage in perspective taking and attribute mental states to the self and to others to successfully predict a person's behavior. Children participated in 2 false belief tasks, a location change task and an unexpected contents task. The "Sally-Anne" location change task (Baron-Cohen et al., 1985) presented children with the following scenario: "Sally puts her marble in the basket. Then she goes out for a walk. While she is gone, Anne takes the marble out of the basket and puts it in the box. When Sally comes back, where will she look for her marble (belief question)? Where is the marble really (control memory question)? Where was the marble in the beginning (control memory question)?" Children only received a point for correctly answering the belief question if they also answered the memory questions correctly.

(2) Unexpected contents task. A variation of the "Smarties" unexpected contents task was also be administered (Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987). The original "Smarties" task showed participants that a "Smarties" container had

pencils instead of candy and asked children what someone who has not seen what was inside the "Smarties" container would think was inside the container. In this variation of the task children were shown a crayon box and were asked what they think is inside, then they were shown that candles are inside the box (Atance, Bernstein, & Meltzoff, 2010). After displaying the content inside the box, children were asked the following questions: "When you first saw this box, before we opened it, what did you think was inside (belief question)? When Sammy (a stuffed animal), walks in and sees the box, what will Sammy think is inside the box (belief question)? What is really inside the box (control memory question)?" Children only received points for correctly answering the belief questions if they also answered the memory question correctly. Children were able to earn up to 2 points on this task. These false belief tasks usually show a reliable sex difference in normative samples, with boys performing more poorly than girls (Walker, 2005). Administration of both TOM tasks took about 5 minutes and total scores ranged from 0 to 3 points.

Eye-tracking Tasks: (1) Preferential looking task. Children were presented with 3 park scenes displaying the same park objects (e.g., fountain, airplane, swing-set) and people on a computer monitor for 5 seconds each. The background for all park scenes remained constant, while the location of the people and objects was different for all scenes. An Applied Science Laboratory D6 model eye-tracking system was used to determine whether children spent more time looking at people or objects. Although theses stimuli has never been used before, studies using similar stimuli have shown a

reliable sex difference in normative samples, with boys preferring to look at objects and girls preferring to look at people (Connellan et al., 2000).

(2) Emotion recognition task. Second, Children were asked to identify facial emotions of 20 pictures from the National Institute of Mental Health Child Emotional Faces Picture Set (NIMH-ChEFS) (Egger et al., 2011). The NIMH-ChEFS contains 482 colored pictures of children between the ages of 10 to 17-years who are expressing 5 different facial emotions (afraid, angry, happy, neutral, and sad), under 2 gaze conditions (direct and averted). Male and female pictures with the highest goodness score (i.e., highest accuracy and intensity ratings) in each emotion category and gaze condition were used as stimuli for this project. As children viewed the 20 pictures on a computer monitor, they were asked to identify which of the 5 emotions (scared, mad, happy, neutral/ok, or sad) each picture is displaying. Eye-tracking data was used to examine which region of the face (e.g., eyes, mouth) children spent more time looking at during the emotion recognition task. Similar emotion recognition tasks have shown a reliable sex difference in normative samples, with boys performing worse than girls (Golan et al., 2006). Eye-tracking tasks took approximately 10 to 15 minutes to complete.

2.3 Procedure

Parents were asked to complete the questionnaires while their children participated in the experimental tasks. Children first completed the EOWPVT and the mental rotation task. Then they were allowed to take a 5 minutes break outside of the testing room. Next, they completed the eye-tracking tasks and had their hand

photocopied. At this time children were allowed to take another 5 minute break. Finally, they completed the PPPSI and false belief tasks. The tasks were grouped in this manner to ensure that the three sessions would take approximately the same amount of time to complete (20 minutes). Additionally, the eye-tracking tasks were administered in the middle of the testing because these tasks required the child to move into the adjacent room, allowing them to change their scenery. Children had coloring books and crayons available to them during their breaks and received a small prize for participating.

3. RESULTS

Preliminary data analyses were conducted to ensure all data were normally distributed and outliers were identified using the outlier labeling rule (Hoaglin & Iglewicz, 1987; Hoaglin, Iglewicz, & Tukey, 1986). One outlier was identified in the children's autism quotient and mental rotation task total score. These outliers were replaced with the sum of the mean and standard deviation for their respective scales.

3.1 Sex differences

A one-way between subjects analysis of variance (ANOVA) was conducted to determine if boys and girls were significantly different on any of our variables of interest (see Table 1). No significant differences emerged on any measure, except for the Playmate and Play Style Preferences Structured Interview (PPPSI). On the PPPSI girls rated themselves as preferring more feminine playmates (girls: 10.43 ± 4.26 *vs.* boys: 4.09 ± 4.74, $d = 1.41$) and play styles (girls: 8.48 ± 2.73 *vs.* boys: 5.04 ± 2.40, $d = 1.34$) than boys, $F(1, 44) = 22.82, 20.54$, $p < .001$. Although non-significant, it is important to note that the means for postnatal testosterone levels, Child Behavior Checklist total problems, autism quotients, percent of time participants spent looking at people while viewing the park scenes, and mental rotation accuracy scores were not in the expected direction. Interestingly, all of the means for variables generated from the experimental tasks were in the expected direction except for the mental rotation accuracy scores, where girls scored slightly better than boys.

Table 1.
Sex Differences across Study Variables.

Variable	Boys (*M*, *SD*)	Girls (*M*, *SD*)	*p*	*d*
Hormones				
Salivary T levels (pg/ml)	39.39 (10.53)	41.80 (13.86)	.49	.20
2D:4D ratios (mm)	.93 (.03)	.95 (.03)	.15	.67
Questionnaires				
CBCL total problems	29.37 (19.21)	31.13 (20.99)	.72	.09
AQ-C	133.69 (12.20)	135.08 (12.08)	.64	.11
EQ-C	30.74 (9.39)	31.97 (8.96)	.59	.13
SQ-C	24.29 (8.89)	21.84 (7.33)	.23	.30
Experimental tasks				
EOWPVT	66.29 (21.31)	71.70 (16.34)	.34	.28
PPPSI playmate	4.09 (4.74)	10.43 (4.26)	.00**	1.41
PPPSI play style	5.04 (2.40)	8.48 (2.73)	.00**	1.34
MRT accuracy	12.30 (4.19)	12.96 (3.98)	.59	.16
MRT total time (ms)	81.68 (24.54)	84.95 (49.15)	.78	.08
ERT	14.96 (3.94)	15.48 (3.86)	.65	.13
False belief measures	3.70 (1.11)	3.96 (1.07)	.42	.24
Eye tracking (LZ % time)				
Eyes	31.19 (17.57)	32.37 (17.72)	.86	.07
Mouth	27.56 (14.92)	25.85 (19.12)	.77	.10
People	9.57 (3.68)	8.01 (4.09)	.26	.40

Note: T, Testosterone; CBCL, Child Behavior Checklist; AQ-C, Autism Spectrum Quotient-Children's Version; EQ-C, Empathizing Quotient-Children's Version; SQ-C, Systemizing Quotient-Children's Version; EOQPVT, Expressive One-Word Picture Vocabulary Test; PPPSI, Playmate and Play Style Preferences Structured Interview; MRT, Mental Rotation Task; ERT, Emotion Recognition Task; LZ, Look Zone.
** $p < .001$.

In addition to examining sex differences using the total false belief score obtained from adding results from the two false belief measures (presented in table 1), two chi-square tests of independence were also conducted to test whether there was a relationship between a child's sex (girl vs. boy) and each of the false belief measures (theory of mind vs. no theory of mind). For the "Sally-Anne" location change task there was a significant relationship between theory of mind and a child's sex, x^2 (1, N = 46) = 4.26, p < .05. Results indicated that most girls passed this false belief task (65% of girls passed this task) while most boys did not (35% of boys passed this task). For the unexpected contents task there was no significant relationship between theory of mind and a child's sex, x^2 (1, N = 46) = .79, p > .05 (48% of girls and 61% of boys passed this task).

3.2 Correlations

Pearson Product Moment correlations were calculated between hormone markers and behavioral measures (see Table 2). Higher (more feminine) 2D:4D ratios in both sexes at 5-years were correlated with higher empathizing scores, r (45) = .35, p = < .05, higher expressive vocabulary scores, r (45) = .35, p = < .05, longer time completing the mental rotation task, r (45) = .37, p = < .05, higher scores on the emotion recognition task, r (45) = .46, p = < .01, and more time spent looking at the eye region of faces 5-years, r (34) = .41, p = < .05. Higher (more masculine) salivary testosterone levels in both sexes at three-months of age were correlated with less time spent looking at people instead of objects at 5-years, r (26) = -.42, p = < .05. Unexpectedly, higher salivary

Table 2.
Pearson Product Moment Correlations between Hormones and Behaviors

	T levels (pg/ml)	2D:4D ratios (mm)
Questionnaires		
CBCL total problems	-.20	-.08
AQ-C	.07	.17
EQ-C	.33*	.35*
SQ-C	.08	.13
Experimental tasks		
EOWPVT	.07	.35*
PPPSI playmate	.17	-.01
PPPSI play style	.36*	.18
MRT accuracy	-.02	-.08
MRT total time (ms)	.09	.37*
ERT	.05	.46**
False belief measures	.14	-.14
Eye tracking (LZ % time)		
Eyes	-.25	.41*
Mouth	-.07	-.10
People	-.42*	.15

Note: T, Testosterone; CBCL, Child Behavior Checklist; AQ-C, Autism Spectrum Quotient-Children's Version; EQ-C, Empathizing Quotient-Children's Version; SQ-C, Systemizing Quotient-Children's Version; EOQPVT, Expressive One-Word Picture Vocabulary Test; PPPSI, Playmate and Play Style Preferences Structured Interview; MRT, Mental Rotation Task; ERT, Emotion Recognition Task; LZ, Look Zone.
* $p < .05$. ** $p < .01$.

Table 3.
Pearson Product Moment Correlations between Empathizing/Systemizing Scores and Other Measures

	Empathizing Quotient	Systemizing Quotient
Questionnaires		
CBCL total problems	-.54**	.09
AQ-C	.45**	-.42**
Experimental tasks		
EOWPVT	.41**	.10
PPPSI playmate	.09	.13
PPPSI play style	.12	.20
MRT accuracy	.29	.04
MRT total time (ms)	.13	.03
ERT	.31*	.29
False belief measures	.12	.00
Eye tracking (LZ % time)		
Eyes	.15	.25
Mouth	.03	-.16
People	.25	.26

Note: T, Testosterone; CBCL, Child Behavior Checklist; AQ-C, Autism Spectrum Quotient-Children's Version; EQ-C, Empathizing Quotient-Children's Version; SQ-C, Systemizing Quotient-Children's Version; EOQPVT, Expressive One-Word Picture Vocabulary Test; PPPSI, Playmate and Play Style Preferences Structured Interview; MRT, Mental Rotation Task; ERT, Emotion Recognition Task; LZ, Look Zone.
* $p < .05$. ** $p < .01$.

testosterone levels were also correlated with higher empathizing scores, r (50) = .33, p = < .05, and more feminine play styles, r (37) = .36, p = < .05.

In addition, Pearson r correlation coefficients were calculated to explore the relationship between empathizing/systemizing scores and other measures collected in this study (see Table 3). Empathizing and systemizing scores were found to relate positively to one another, r (66) = .26, p = < .05. Empathizing scores also positively correlated with autism quotients, r (66) = .46, p = < .01, expressive vocabulary scores, r (47) = .41, p = < .01, and emotion recognition scores, r (46) = .31, p = < .05, and negatively correlated with CBCL total problem scores, r (66) = -.54, p = < .01. Additionally, systemizing scores correlated negatively with autism quotients, r (66) = -.42, p = < .01.

Lastly, the Pearson r correlation coefficient was calculated between children and parents' autism quotient scores. Results indicated that there was a positive relationship between children and parents' autism quotient scores, r (63) = .50, p = < .05, such that parents who endorsed more autism traits also reported higher incidence of autism traits in their children.

3.3 Regressions

Simultaneous multiple regressions were conducted to assess whether hormone levels would be predictive of variables that they were previously correlated with, even after controlling for a child's sex and age (see Table 4). A total of 7 regressions were

Table 4.
Summary of Simultaneous Multiple Regressions for All Infants

	B	*SE*	*β*
EQ-C[a]			
Sex	-2.41	2.19	-.17
Age (days)	.02	.01	.49**
2D:4D ratios (mm)	7.05	40.55	.03
Testosterone levels (pg/ml)	.13	.10	.19
EOWPVT[b]			
Sex	4.16	5.01	.11
Age (days)	.06	.01	.73**
2D:4D ratios (mm)	29.46	92.75	.04
T levels (pg/ml)	.15	.22	.09
PPPSI play style[c]			
Sex	-4.08	.89	-.63**
Age (days)	-.00	.00	-.13
2D:4D ratios (mm)	2.50	16.43	.02
T levels (pg/ml)	.06	.04	.19
MRT total time (ms)[d]			
Sex	7.17	13.52	.09
Age (days)	-.03	.03	-.17
2D:4D ratios (mm)	882.31	250.64	.57**
T levels (pg/ml)	.35	.60	.09
ERT[e]			
Sex	1.55	1.13	.21
Age (days)	.01	.00	.51**
2D:4D ratios (mm)	29.09	20.89	.21
T levels (pg/ml)	.03	.05	.09
Eyes (LZ % time)[f]			
Sex	-5.84	6.77	-.19
Age (days)	.01	.01	.18
2D:4D ratios (mm)	97.92	122.29	.17
T levels (pg/ml)	-.42	.28	-.32
People (LZ % time)[g]			
Sex	.35	1.43	.05
Age (days)	.01	.00	.40*
2D:4D ratios (mm)	-11.71	28.87	-.07
T levels (pg/ml)	-.13	.06	-.43*

[a] $R^2 = .36$, $F\ (4, 31) = 4.30$, $p < .05$.
[b] $R^2 = .55$, $F\ (4, 31) = 9.46$, $p < .05$.
[c] $R^2 = .49$, $F\ (4, 31) = 7.56$, $p < .05$.
[d] $R^2 = .29$, $F\ (4, 31) = 3.20$, $p < .05$.
[e] $R^2 = .37$, $F\ (4, 31) = 4.58$, $p < .05$.
[f] $R^2 = .17$, $F\ (4, 21) = 1.14$, $p > .05$.
[g] $R^2 = .33$, $F\ (4, 21) = 2.63$, $p > .05$.

Note: T, Testosterone; EQ-C, Empathizing Quotient-Children's Version; EOQPVT, Expressive One-Word Picture Vocabulary Test; PPPSI, Playmate and Play Style Preferences Structured Interview; MRT, Mental Rotation Task; ERT, Emotion Recognition Task; LZ, Look Zone.
* p < .05. ** p < .01.

conducted using sex (dummy-coded 0 = female, 1 = male), age, 2D:4D ratios, and testosterone levels as independent variables. These findings are discussed below.

The overall regression used to predict empathizing scores was significant, F (4, 31) = 4.30, $p < .05$, $R^2 = .36$. Of the predictors investigated, age was a significant predictor of empathizing scores, $\beta = .49$, t (31) = 3.24, $p < .05$. That is, after controlling for sex, 2D:4D ratios, and testosterone levels, a one-point increase in age (one day older) resulted in a .02 point predicted increase in empathizing scores. Age accounted for 22% of the variance in empathizing scores.

The overall regression used to predict expressive vocabulary scores was also significant, F (4, 31) = 9.46, $p < .05$, $R^2 = .55$. Of the predictors investigated, age was a significant predictor of expressive vocabulary scores, $\beta = .73$, t (31) = 5.76, $p < .05$. That is, after controlling for sex, 2D:4D ratios, and testosterone levels, a one-point increase in age (one day older) resulted in a .06 point predicted increase in expressive vocabulary scores. Age accounted for 48% of the variance in expressive vocabulary scores.

Similarly, the overall regression used to predict play styles was significant, F (4, 31) = 7.56, $p < .05$, $R^2 = .49$. Of the predictors investigated, sex was a significant predictor of play styles, $\beta = -.63$, t (31) = -4.60, $p < .05$, such that females preferred more female typical play styles. Sex accounted for 35% of the variance in play styles.

The overall regression used to predict completion time of the mental rotation task was also significant, F (4, 31) = 3.20, $p < .05$, $R^2 = .29$. Of the predictors investigated, 2D:4D ratios was a significant predictor of mental rotation completion time, $\beta = .57$, t (31) = 3.52, $p < .05$. That is, after controlling for sex, age, and testosterone levels, a one-

point increase in 2D:4D ratios resulted in a 882.31 point predicted increase in the time (measured in milliseconds) used to complete the mental rotation task. In other words, more feminine 2D:4D ratios were predictive of slower completion times. In total, 2D:4D ratios accounted for 28% of the variance in mental rotation completion time.

Likewise, the overall regression used to predict scores on the emotion recognition task was significant, F (4, 31) = 4.58, $p < .05$, $R^2 = .37$. Of the predictors investigated, age was a significant predictor of emotion recognition scores, $\beta = .51$, t (31) = 3.40, $p < .05$. That is, after controlling for sex, 2D:4D ratios, and testosterone levels, a one-point increase in age (one day older) resulted in a .01 point predicted increase in emotion recognition scores. Age accounted for 23% of the variance in emotion recognition scores.

The regression used to predict the percent of time that participants' spent looking at the eye region of faces was not significant (F (4, 22) = 1.14, $p > .05$, $R^2 = .17$) and no main effects emerged.

The overall regression used to predict the percent of time spent looking at people compared to objects was not significant, F (4, 21) = 2.63, $p > .05$, $R^2 = .33$. Of the predictors investigated, both age ($\beta = .40$, t (21) = 2.21, $p < .05$) and testosterone levels ($\beta = -.43$, t (21) = -2.18, $p < .05$) were significant predictors of the time spent looking at people instead of objects. That is, after controlling for sex, 2D:4D ratios, and testosterone levels, age accounted for 16% of the variance in the percent of time spent looking at people, such that a one-point increase in age (one day older) resulted in a .01 point predicted increase in the percent of time spent looking at people. Similarly, after

controlling for sex, age, and 2D:4D ratios, testosterone levels accounted for 15% of the variance in percent of time spent looking at people relative to objects, such that higher testosterone levels were predictive of less time spent looking at people.

3.4 Dependent samples t-tests

Various dependent-samples t tests were conducted to test whether specific dimensions of the emotion recognition task influenced the scores obtained on this measure. First, the stimuli for the emotion recognition task were divided by sex and the mean score for male faces was compared to the mean score for female faces. Results indicated that participants were equally able to correctly identify emotions when they viewed the face of a male ($M = 7.74$, $SD = 2.12$) compared to when they viewed the face of a female ($M = 7.48$, $SD = 1.99$), $t\ (45) = 1.26$, $p > .05$. The stimuli for the emotion recognition task were also divided by direction of gaze and the mean score for direct faces (faces of children who looked directly at the camera when their photograph was taken) was compared to the mean score for averted faces (faces of children who did not look directly at the camera when their photograph was taken). Overall, participants obtained higher scores on the emotion recognition task when they viewed stimuli with averted eye-contact ($M = 7.83$, $SD = 2.11$) compared to when they viewed stimuli with direct eye-contact ($M = 7.39$, $SD = 1.94$), $t\ (45) = -2.41$, $p < .05$. Results also indicated that participants spent the same percent of time viewing the eye region of the faces ($M =$ 31.78%, $SD = 17.40\%$) and the mouth region of the faces ($M = 26.63\%$, $SD = 17.10\%$), t $(34) = 1.45$, $p > .05$.

A one-way within subjects analysis of variance (ANOVA) was conducted to determine if participants were better able to identify at least one of the five emotions presented in the face stimuli during the emotion recognition task. The scores for the children were significantly different across the different emotions, Greenhouse-Geisser adjusted F (3.14, 141.28) = 13.12, p $< .05$. Dependent samples t tests were conducted to assess which of the emotions differed from one another, with each test conducted at an alpha level of .005. The results indicated that participants were more accurate at identifying mad feelings (M = 3.41, SD = .98) than scared feelings (M = 2.67, SD = 1.37), t (45) = -3.65, $p < .005$, and happy feelings (M = 7.83, SD = 2.11) than scared feelings, t (45) = -4.24, $p < .005$. Participants were also more accurate at identifying mad feelings than ok/neutral feelings (M = 2.26, SD = 1.61), t (45) = -4.60, $p < .005$, happy feelings than ok/neutral feelings, t (45) = -5.41, $p < .005$, and sad feelings (M = 3.26, SD = 95) than ok/neutral feelings, t (45) = -3.87, $p < .005$. There was not a significant difference in accuracy of emotion recognition between sad and scared stimuli, t (45) = -2.66, $p > .005$, sad and mad stimuli, t (45) = .88 $p > .005$, sad and happy stimuli, t (45) = 2.23, $p > .005$, scared and ok/neutral stimuli, t (45) = 1.55, $p > .005$, or mad and happy stimuli, t (45) = -1.32, $p > .005$.

Pearson moment correlations were conducted to test whether testosterone levels were related to the recognition of specific emotions. Results indicated that postnatal testosterone was not associated with any of the five emotions. In contrast, higher 2D:4D ratios (more feminine) were related to better identification of scared, r (45) = .42, $p = <$ .05, and ok/neutral emotions, r (45) = .34, $p = < .05$.

Several dependent-samples t tests were also conducted to explore the eye-tracking results obtained while participants viewed 3 park scenes. First, the mean percent of time spent looking at people was compared to the mean percent of time spent looking at objects. Results indicated that participants spend more time looking at objects (M = 63.88, SD = 17.09) than people (M = 8.70, SD = 3.94), t (33) = -15.43, $p < .05$. Participants also spent more time looking at males (M = 7.73, SD = 6.51) than females (M = 5.35, SD = 3.66), t (33) = 2.16, $p < .05$, and males carrying a baby (M = 11.42, SD = 7.06) compared to females carrying a baby (M = 7.42, SD = 5.29), t (33) = 3.79, $p < .05$. Results also indicated that participants spent significantly more time looking at adults carrying a baby (M = 9.42, SD = 5.42) than adults without a baby (M = 6.54, SD = 4.20), t (33) = -4.45, $p < .05$.

4. DISCUSSION

4.1 Postnatal testosterone and sex-typed behavior

The main purpose of this study was to increase our understanding of the relationship between the postnatal surge of testosterone and later patterns of behavior by examining whether postnatal testosterone levels collected at 3-months of age would be predictive of sex-typed behaviors in early childhood. As hypothesized, after controlling for levels of prenatal androgen exposure (using 2D:4D ratios), greater postnatal testosterone levels in early infancy were predictive of less time spent looking at people when presented with a complex park scene with people and objects (e.g., swing set, spring rider, water fountain, airplane, gazebo, picnic table, trash can, flower bed) embedded in it. To our knowledge, this is the first study to document a relationship between the postnatal surge of testosterone and sex-linked behavior past the second year of life. Our finding suggests that the postnatal surge of testosterone that occurs in early infancy could have organizational effects on behavior. However, independent replication of these findings and future research establishing a relationship between postnatal testosterone and behavior beyond early childhood is required to establish the human postnatal period as a separate critical period for the development of later behavior.

Previous studies have suggested behavioral outcomes of the postnatal endocrine surge at earlier times in development. For example, higher postnatal testosterone levels predicted stronger male-typical visual preferences and temperament traits at 3-months (Alexander & Saenz, 2011; Alexander, Wilcox, & Farmer, 2009), as well as higher

autism spectrum traits and weaker language development at the age of 2-years (Saenz & Alexander, 2013b). Similarly, others have reported higher postnatal androgen levels influence the development of masculine toy preferences and delay maturation of brain structures necessary for language functioning (Friederici et al., 2008; Lamminmäki et al., 2012).

One suggestion based on findings from early primate research and recent studies in human behavior is that the postnatal endocrine surge may be necessary for the developmental programming of male social relations (Alexander, 2014). Since interest in people is generally considered a female trait necessary to facilitate female's advantage for forming social relations (Baron-Cohen, 2010), the present research suggests greater postnatal testosterone levels result in less visual interest in people (female trait) and more visual interest in objects (male trait). Therefore, it is possible that a slight increase in postnatal testosterone levels may lead to small deficits in socialization that fall within a normal range of variation for typically developing children, while a large increase in postnatal testosterone levels may lead to large deficits in socialization similar to what is observed in children with an ASD and other disorders characterized by deficits in socialization. For that reason, it may be useful to examine social functioning in children born prematurely and who consequently undergo an exaggerated rise in testosterone during the postnatal surge (Tapanainen, Koivisto, Vihko, & Huhtaniemi, 1981) as well as for children exposed in early life to substances in food and plastics (e.g., bisphenol A and phthalates) that disrupt normal hormonal processes (de Cock, Maas, & van de Bor, 2012; Jurewicz & Hanke, 2011).

It may also be the case that our finding of a relationship between postnatal testosterone levels and visual interest in people is spurious. Higher postnatal testosterone levels were not predictive of other more male-typical behaviors (i.e., lower empathy quotients, lower expressive language scores, lower mental rotation reaction times, less time spent looking at eye region of faces, lower facial emotion recognition scores, greater preference for masculine play styles), and contrary to our expectation were positively correlated with some female-typical behaviors (i.e., more feminine play styles and higher empathizing scores). Alternatively, despite broad generalizations (i.e., postnatal testosterone masculinizes all gender-typed behaviors), these findings may illustrate the complex relation that exists between hormones and behavior and show the utility of measuring different components of gender-linked behavior. A comprehensive theoretical model that can account for human hormone-behavior effects will require collaboration across laboratories and disciplines. However, the outcome of this effort may result in a better understanding of the mini-puberty of infancy and its implications for both typical and atypical social development.

Previous studies conducted in our laboratory that have used this same cohort of participants at former phases of development have found no relationship between postnatal testosterone levels and toy preferences. For instance, when these children were 4-months-old, we found no significant relationship between postnatal testosterone levels and the amount of time that they looked at a ball compared to a doll, measured via an eye-tracking preferential looking task (Alexander, Wilcox, & Woods, 2009). Similarly, when these children were 2-years-old, we found no significant relationship between

postnatal testosterone levels and the percent of time that they spent playing with female-typical toys (Alexander & Saenz, 2012). In contrast, the present study found a significant relationship between postnatal testosterone levels and feminine play styles at the age of 5-years. However, after controlling for a child's sex, the relationship between postnatal testosterone and play styles was no longer significant.

Our positive association between postnatal testosterone levels and feminine play styles is also interesting given that prior research examining the relationship between postnatal testosterone levels and toy preferences outside of our laboratory has also yielded mixed findings. For example, one study investigated the relationship between postnatal testosterone levels in infancy and the percentage of time that toddlers played with various toys, and found that postnatal testosterone levels in boys were negatively correlated with the amount of time that they played with a doll and a truck (Lamminmäki et al., 2012). Together, our findings along with these findings indicate that it remains unclear whether postnatal testosterone influences the development of gender-linked toys and indicate a need to further investigate this relationship in the future.

Our results also indicated that higher postnatal testosterone levels were significantly correlated with higher empathizing scores. We expected to find an inverse relationship between these two measures since the ability to empathize, that is understand others' feelings and thoughts, is believed to be better developed in females compared to males (Baron-Cohen, 2010). We selected to use the empathizing quotient measure for this study because it purported to show a medium effect size (Cohen's d = .76) between boys and girls (Chapman et al., 2006). However in our sample, girls

obtained slightly higher empathizing scores than boys, but this difference was small and non-significant. Perhaps our small sample size contributed to these findings.

4.2 Prenatal testosterone and sex-typed behavior

We also examined the relationship between 2D:4D ratios and sex-typed behavior in early childhood. Previous research has reported large sex differences in performance of mental rotation tasks in adults (Peters, Manning, & Reimers, 2007), with males outperforming females. However, a recent review indicates that these sex differences in performance of mental rotation tasks are rarely found before the age of 9 (Frick, Möhring, & Newcombe, in press). Consistent with this finding, males and females in our study performed similarly on the mental rotation task. The source of this sex difference is debatable, but studies reporting sex differences in infants as young as 3-months (Moore & Johnson, 2008; Quinn & Liben, 2008) and findings that females with abnormally high levels of prenatal testosterone exposure due to having an opposite-sex twin or Congenital Adrenal Hyperplasia perform better than controls (Berenbaum et al., 2012; Puts, McDaniel, Jordan, & Breedlove, 2008; van Anders, Vernon, & Wilbur, 2006) have suggested a role for biological factors. Contrary to the hypothesis that prenatal testosterone exposure masculinizes visual-spatial ability, we found no relationship between 2D:4D ratios and mental rotation accuracy scores. However, our finding is consistent with some other prior research, indicating that mental rotation ability in children is unrelated to fetal testosterone levels measured in amniotic fluid collected during amniocentesis (Auyeung, Knickmeyer, et al., 2012). In addition, similar

studies conducted in adults yield mixed findings, with some indicating that 2D:4D ratios are unrelated to mental rotation ability (Anders & Hampson, 2005) and others indicating that smaller 2D:4D ratios are related to better performance on mental rotation tasks (Peters et al., 2007). A meta-analysis combining child and adult studies, concluded that there is no reliable association between 2D:4D ratios and mental rotation ability (Puts et al., 2008), consistent with the results of this investigation.

Yet, individuals who were exposed to higher levels of prenatal testosterone as indicated by digit ratios were faster at completing the mental rotation task. Previous research examining this association has yielded varied results. For example, one study found that higher fetal testosterone levels measured in amniotic fluid were associated with a faster mental rotation response time among 7-year-old girls but not boys (Grimshaw et al., 1995), while a similar study found that fetal testosterone levels measured in amniotic fluid were not related to mental rotation response times of 7-year-old boys or girls (Auyeung, Knickmeyer, et al., 2012). These disparate findings could be a result of differences in the methodological design of these studies (e.g., measurement of prenatal testosterone, presentation of stimuli, number of trials in mental rotation task) and indicate a need for more systematic analyses of testosterone and the factors that contribute to performance in children.

Our results also indicated that more feminine 2D:4D ratios (indicative of less prenatal testosterone exposure) were related to higher scores on the emotion recognition task. This finding is consistent with previous research indicating a negative relationship between testosterone levels measured in amniotic fluid and higher scores on an emotion

recognition task requiring children to determine a person's emotional state while only looking at the eye region of a face ("Reading the Mind in the Eyes") (Chapman et al., 2006). Furthermore, children with autism tend to have more masculine 2D:4D ratios (indicative of greater prenatal testosterone) and also score lower on emotion recognition tasks (Bal et al., 2010; P. Krajmer, Spajdel, Kubranska, & Ostatnikova, 2011). Together, these studies support the idea that 2D:4D ratios are sensitive to the factors that influence the social-emotional development that occurs in early childhood.

Additionally, we found a positive association between 2D:4D ratios and scores on the Expressive One-Word Picture Vocabulary Test, such that higher 2D:4D ratios (indicative of less prenatal testosterone exposure) were related to higher vocabulary scores. These findings are consistent with previous findings from our laboratory indicating that higher 2D:4D ratios (indicative of less prenatal testosterone exposure) are predictive of larger vocabularies among 2-year-old toddlers (Saenz & Alexander, 2013b). Likewise, other researchers have noted an inverse relationship between prenatal testosterone measured in amniotic fluid and vocabulary size among 2-year-old toddlers (Lutchmaya et al., 2002b). Moreover, lower 2D:4D ratios (indicative of more prenatal testosterone exposure) have also been related to more articulation problems among 3-year-old children (Albores-Gallo, Fernandez-Guasti, Hernandez-Guzman, & List-Hilton, 2009), while boys with higher prenatal testosterone levels measured in umbilical cord blood have been found to be at greater risk for developing a language disorder within the first three years of life (Andrew J. O. Whitehouse et al., 2012). Furthermore, greater fetal androgen exposure has been associated with greater pragmatic language difficulties

among 10-year-old girls (Andrew J.O. Whitehouse et al., 2010), and in adults, higher 2D:4D ratios (indicative of less prenatal testosterone exposure) have been positively correlated with verbal intelligence (Luxen & Buunk, 2005). This research compellingly supports the idea that prenatal testosterone is influential in shaping our verbal development throughout the lifespan.

Previous research has documented that typically developing females score higher on the Empathizing Quotients-Children's Version questionnaire than typically developing males (Auyeung, Wheelwright, et al., 2009). Although not statistically significant, our results yielded a similar pattern with females obtaining higher empathizing scores than males. Our results also indicated a positive correlation between 2D:4D ratios and empathizing quotients, such that children with more feminine 2D:4D ratios (indicative of less prenatal testosterone exposure) obtained higher empathizing scores. Theoretically this makes sense since the ability to empathize is generally considered a trait that females are better at than males, and therefore we would expect more feminine 2D:4D ratios to result in more feminine traits, including empathy. Our finding is consistent with research indicating an inverse relationship between fetal testosterone measured in amniotic fluid and empathizing quotients in 6- to 9-year-old children (Chapman et al., 2006). Yet, other studies using the adult version of the Empathizing Quotient questionnaire have produced mixed results; with some studies reporting outcomes similar to ours (Wakabayashi & Nakazawa, 2010), and others failing to find a relationship between 2D:4D ratios and empathizing quotients (Johannes Hönekopp, 2012; John T. Manning, Baron-Cohen, Wheelwright, & Fink, 2010). It is

likely that methodological differences contribute to these varied findings in adult samples, since the studies reporting no association between these variables have primarily consisted of online studies that rely on participants to measure and report their own finger lengths.

Finally, our results revealed a positive correlation between 2D:4D ratios and the amount of time that children spent looking at the eye portion of a face while completing an emotion recognition task. In other words, children with more feminine 2D:4D ratios (indicative of less prenatal testosterone exposure) spent more time looking at the eye region of faces. This result is consistent with previous findings from our laboratory indicating that more feminine 2D:4D ratios (indicative of less prenatal testosterone exposure) are predictive of longer durations of eye contact and more frequent eye contact between toddler-parent dyads during an 8-minute, unstructured, play interaction (Saenz & Alexander, 2013a). This finding is also consistent with research conducted outside of our laboratory indicating that higher fetal testosterone exposure, measured in amniotic fluid, is related to less eye contact between infant-parent dyads and lower scores on the "Reading the Mind in the Eyes" task (an emotion recognition task that only presents children with the eye region of faces) (Chapman et al., 2006; Lutchmaya et al., 2002a).

Attention to the eye region of faces is important since it has been linked to greater competency in processing social-emotional cues during social interactions and facilitates emotion recognition in others. For example, research has found that more gaze to the eye region of faces, during an emotion recognition task, is related to better

emotion recognition scores in children with an Autism Spectrum Disorder (Karabekiroglu et al.) (Bal et al., 2010). Research has also identified different scanning patterns between typically developing children and children with an ASD. One study revealed that while scanning patterns of typically developing children remain similar between 2- and 4-years of age, children with an ASD tend to spend less time looking at the eyes, mouth, and nose region of faces at 4-years than at 2-years (Chawarska & Shic, 2009). Similarly, more recent research indicates that children who eventually go on to develop an ASD engage in the same amount of eye contact as do typically developing newborns, but in contrast to their typically developing peers those who go on to develop an ASD experience a significant decline in eye contact from 2- to 6-months of age (Jones & Klin, 2013). The use of eye tracking technology has made it possible to detect characteristics of ASD at a younger age, enabling earlier diagnosis and treatment.

Overall, more masculinized 2D:4D ratios (indicative of more prenatal testosterone exposure) were related to more masculine behaviors (i.e., lower empathizing scores, lower expressive vocabulary scores, shorter time required to complete the mental rotation task, lower emotion recognition scores, and less time spent looking at the eye region of faces). Although there has been debate concerning the validity of digit ratios as a proxy measure of prenatal androgen exposure (Berenbaum et al., 2009), these results provide further support for the use of 2D:4D ratios as a proxy measure of prenatal testosterone exposure and the masculinizing effect that early prenatal androgens exert on the development of subsequent sex-typed behaviors. These findings also support the extreme male brain theory of autism, which suggests that ASD may be a result of very

masculinized traits partly caused by high levels of prenatal androgen exposure (Baron-Cohen, 2010).

4.3 The extreme male brain theory of autism

The extreme male brain theory of autism posits that females are generally better able to detect and understand others' emotional state (empathizing), while males are generally better able to construct systems and detect patterns (systemizing). Although these traits are expected to vary from person to person, it is speculated that the ability to empathize is positively related to other abilities that females outperform males in and that this profile is what constitutes a female brain, while systemizing is believed to be related to other abilities that males outperform females in and this profile is what constitutes a male brain (Baron-Cohen, 2002). According to this theory, the majority of people are equally good at empathizing and systemizing, while some are slightly better at one ability than the other, and a small group of people are extremely better at one ability than the other. People with an ASD and others who lack social skills but have strengths in visual spatial abilities make up the small group of people who are extremely better at systemizing than empathizing. Previously, a small non-significant, negative relationship has been reported between empathizing and systemizing scores in typically developing children and in children with an ASD (Auyeung, Wheelwright, et al., 2009). More recently, the strength of the inverse relationship between systemizing and empathizing scores has been noted to vary by genetic vulnerability to ASD, such that the relationship is strongest in individuals diagnosed with ASD, followed by first-degree

relatives, and a control group (Grove, Baillie, Allison, Baron-Cohen, & Hoekstra, 2013). In contrast to previous research, we found a small positive relationship between empathizing and systemizing scores. One possible explanation for our positive relationship could be due to the fact that we combined the questions from the empathizing and systemizing quotient questionnaires and presented the parents with one single questionnaire. This makes it likely that the parents either presented their child in a very favorable or negative light throughout the entirety of the questionnaire, regardless of what the questions were asking (e.g., "My child likes to look after other people" vs. "My child remembers large amounts of information about a topic that interests them").

To further test Baron-Cohen's theory predicting that empathizing/systemizing scores will respectively relate to female/male traits and behaviors, we examined the relationship between empathizing/systemizing scores and results from our other questionnaires and experimental tasks. We found that empathizing quotients were positively related to autism quotient scores, expressive vocabulary scores, and emotion recognition scores, and negatively related to CBCL total problem scores; while systemizing quotients were negatively related to autism quotient scores. These results do not support Baron-Cohen's theory. In particular, the relationship between the empathizing/systemizing quotients and autism quotient was unexpected. We expected that higher autism quotient scores would be related to lower empathizing scores and higher systemizing scores since the diagnostic criteria for ASD includes social-communication impairments, which can be measured with the empathizing questionnaire, and restricted repetitive patterns of behavior, which can be measured with

the systemizing questionnaire. Yet, our results show the opposite relationship. It is difficult to explain our findings, but perhaps our conflicting results are an outcome of our positive association between empathizing and systemizing quotients.

Recently, an internet study conducted on 811 adults with an ASD and 3,906 age-matched typical adults found that control females scored higher on the empathizing quotient, control males scored higher on the systemizing and autism quotients, and females and males with an ASD had extreme male profiles (extremely high scores on systemizing and autistic quotients, below average scores on empathizing quotients), providing further evidence for the extreme male brain theory of autism (Baron-Cohen et al., 2014). Despite strong support for this theory, some researchers have questioned Baron-Cohen's conceptualization and have suggested that measuring empathizing and Machiavellianism makes more sense and fits the data more accurately than measuring empathizing and systemizing (Andrew, Cooke, & Muncer, 2008). Yet other researchers have found that the difficulties with emotion recognition observed in individuals with autism are primarily due to high rates of comorbidities with alexithymia and that these impairments in emotion recognition disappear when ASD and control groups are matched for alexithymia (R. Cook, Brewer, Shah, & Bird, 2013). Unlike the extreme male brain theory of autism, a group of alternative hypotheses (e.g., the intense world hypothesis and the empathy imbalance hypothesis) suggest that individuals with autism do not lack empathy but rather experience heightened levels of emotional sensitivity (Markram, Rinaldi, & Markram, 2007; Smith, 2009).

4.4 Demographic variables

Noteworthy is the fact that although we found associations between hormone markers and behavior, we did not find significant sex differences in most of the measures that we used. Two measures that did yield significant sex differences included the Playmate and Play Style Preferences Structured Interview and one of our false belief measures. Our PPPSI results were consistent with previous findings indicating that females have greater interest in female typical toys and play styles compared to males (Alexander & Hines, 1994). Regarding false belief measures, females were better able to solve the "Sally-Anne" location change task than males, but were just as good as males at solving the unexpected contents task. When we summed points across measures females had more total points than males, but this difference was not significant. Overall these results support prior studies indicating a female advantage for inferring others' thoughts and suggest that location change tasks may be more sensitive to sex differences than unexpected contents tasks (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Happé, 1995).

We also found that a child's age was significantly predictive of higher empathizing scores, expressive vocabulary scores, emotion recognition scores, and percent of time that they spent looking at people while viewing the park scenes. For the most part these results are not surprising given that we expect children to acquire larger vocabularies and more complex forms of thinking as they age. However, our finding indicating that older children spent a longer percent of time looking at people is

interesting. Perhaps this greater interest in people parallels the development of more intricate and interactive play styles that develop later in childhood.

4.5 Limitations and implications

One limitation of our study was its small sample size. Using a statistical calculator we determined that we would need a sample size of approximately 28 participants to obtain significant results from a multiple regression analysis using a probability level of .05, 4 predictors in the model, an anticipated effect size of .35 (large), and a desired statistical power level of .6 (Soper, 2014). Using the same parameters, we would need approximately 58 participants to obtain a medium effect size (f^2 of .15) and 395 participants to detect a small effect size (f^2 of .02). As noted in the results section, five of our multiple regressions included 36 participants and 2 of the multiple regressions included 26 participants. This means that with our current sample size we were only able to detect findings with large effect sizes and it is likely that, due to our limited sample size, we may not have had sufficient power to detect findings with medium or small effect sizes (Cohen, 1992).

Another limitation of this study was that many participants did not complete all components of the experiment. Since this was a longitudinal study that spanned across five years, some original participants were difficult to contact due to various reasons (e.g., new phone number, relocation). Fortunately, many of the participants who relocated agreed to participate remotely. Unfortunately, we were only able to obtain questionnaire data for remote participants and are missing the experimental data. In an

effort to increase our sample size we were also more lenient on the age requirement of participants and extended inclusion criteria from 4- to 5-year-olds to 4- to 7-year-olds.

This research adds to our knowledge of the relationship between testosterone and ASD symptomatology. Previously it has been noted that individuals with high levels of prenatal testosterone, due to Congenital Adrenal Hyperplasia, often score higher on measures of ASD (Knickmeyer et al., 2006). We also know that a subgroup of individuals with an ASD diagnosis have high levels of circulating testosterone (Geier & Geier, 2006a; Ruta, Ingudomnukul, Taylor, Chakrabarti, & Baron-Cohen, 2011), and when these individuals are treated with anti-androgen therapy they experience a decrease in severity of ASD traits (Geier & Geier, 2006b). Similarly, it has been noted in the transgender community that a side effect of hormone treatment for some individuals (specifically testosterone administration that occurs in female-to-male transexuals) results in an increase in ASD symptoms, sometimes pushing people past the threshold needed for an ASD diagnosis (Geier, Kern, King, Sykes, & Geier, 2012). This study provides further support for the relationship between higher prenatal testosterone exposure, measured via 2D:4D ratios, and more ASD symptomatology (e.g., lower empathizing scores, lower vocabulary scores, faster mental rotation times, lower emotion recognition scores, less time spent looking at the eye region of faces while completing an emotion recognition task). Our results also reveal a link between postnatal testosterone levels measured at 3-months of age and decreased visual interest in people compared to objects at the age of 5-years. This greater interest in objects compared to people is characteristic of individuals with ASD (Pierce et al., 2011; Rice, Moriuchi, Jones, &

Klin, 2012). Our study is the first to document a relationship between postnatal testosterone and some ASD behaviors at the age of 5-years.

Knowledge of the interaction between hormones and behavior is central to understanding the development of childhood disorders associated with sex-biased prevalence rates, especially ASD which is described as a set of hyper-masculine characteristics partially caused by high testosterone levels. The existence of a critical period of development in early postnatal life has implications for exposure to environmental chemicals, such as phthalates. Exposure to phthalates during prenatal life has been linked to a decrease in the anogenital distance in male infants (Swan et al., 2005), a decrease in alertness and orientation in female infants (Engel et al., 2009), reduced masculine play in males (Swan et al., 2010), and increase risk of allergy and asthma in male and female children (Jurewicz & Hanke, 2011). More recently, a review reported an association between exposure to endocrine disrupting chemicals (e.g., phthalates and bisphenol A) and increased risk for ASD and ADHD (de Cock et al., 2012). If early postnatal life is indeed another critical period for the development of gender-linked behavior, then exposure to endocrine disruptors during early postnatal life can disrupt the organization of these behaviors.

Future studies should continue to investigate the significance of the postnatal testosterone surge, particularly as it relates to gender-linked behavior in older children, adolescents, and adults. It would also be beneficial to include tasks using eye-tracking technology since this method appears to be more sensitive to individual differences than other experimental tasks. Knowledge can also be gained by investigating the behavioral

effects of the postnatal androgen surge in preterm infants since we know that they experience a higher peak in testosterone levels and recent research has linked higher postnatal testosterone levels in preterm infants to worse health outcomes (Cho, Carlo, Su, & McCormick, 2012).

5. CONCLUSION

The findings of this study add to our understanding of the behavioral impacts that the postnatal androgen surge exerts on sex-linked behaviors. Despite a small sample size, we were able to detect statistically significant relationships between postnatal testosterone levels collected at 3-months and behaviors collected at 5-years. Our results indicate that the postnatal surge of testosterone has some long-term behavioral effects and provides preliminary evidence supporting the hypothesis that the postnatal androgen surge may be another critical period for the organization of sex-linked behavior. To our knowledge, this is the first study to investigate the relationship between postnatal testosterone levels and sex-linked behavior in school-aged children.

A secondary purpose of this study was to investigate the relationship between prenatal testosterone, measured via 2D:4D ratios, and sex-linked behavior. Specifically, we wanted to test Baron-Cohen's extreme male brain theory of autism in a comprehensive manner by collecting quantitative data on various sex-linked behaviors characteristic of ASD symptomatology. On one hand our results showing that lower 2D:4D ratios, indicative of higher prenatal testosterone exposure, were associated with more male-typical behaviors are consistent with the extreme male brain theory of autism. On the other hand, our positive association between empathizing and systemizing scores, as well as their relationship with autism quotient scores was unexpected and incompatible with this theory.

REFERENCES

Achenbach, T. M., & Rescorla, L. A. (2000). Manual for the ASEBA preschool forms and profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Albores-Gallo, L., Fernandez-Guasti, A., Hernandez-Guzman, L., & List-Hilton, C. (2009). 2D:4D finger ratio and language development. *Revista de Neurologia, 48*(11), 577-581.

Alexander, G. M. (2014). Postnatal testosterone concentrations and male social development. *Frontiers in Endocrinology, 5*, 15. doi: 10.3389/fendo.2014.00015

Alexander, G. M., & Hines, M. (1994). Gender labels and play styles: Their relative contribution to children's selection of playmates. *Child Development, 65*(3), 869-879.

Alexander, G. M., & Peterson, B. S. (2001). Sex steroids and human behavior: Implications for developmental psychopathology. *CNS Spectrums, 6*(1), 75.

Alexander, G. M., & Peterson, B. S. (2004). Testing the prenatal hormone hypothesis of tic-related disorders: Gender identity and gender role behavior. *Development and Psychopathology, 16*(2), 407-420.

Alexander, G. M., & Saenz, J. (2011). Postnatal testosterone levels and temperament in early infancy. *Archives of Sexual Behavior, 40*(6), 1287-1292.

Alexander, G. M., & Saenz, J. (2012). Early androgens, activity levels and toy choices of children in the second year of life. *Hormones and Behavior, 62*(4), 500-504. doi: 10.1016/j.yhbeh.2012.08.008

Alexander, G. M., Wilcox, T., & Farmer, M. E. (2009). Hormone-behavior associations in early infancy. *Hormones and Behavior, 56*(5), 498-502.

Alexander, G. M., Wilcox, T., & Woods, R. (2009). Sex differences in infants' visual interest in toys. *Archives of Sexual Behavior, 38*(3), 427-433. doi: 10.1007/s10508-008-9430-1

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text rev. ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anders, S. M. V., & Hampson, E. (2005). Testing the prenatal androgen hypothesis: Measuring digit ratios, sexual orientation, and spatial abilities in adults. *Hormones and Behavior, 47*(1), 92-98.

Andrew, J., Cooke, M., & Muncer, S. J. (2008). The relationship between empathy and Machiavellianism: An alternative to empathizing–systemizing theory. *Personality And Individual Differences, 44*(5), 1203-1211. doi: http://dx.doi.org/10.1016/j.paid.2007.11.014

Arnold, A. P. (1996). Genetically triggered sexual differentiation of brain and behavior. *Hormones and Behavior, 30*(4), 495-505.

Arnold, A. P., & Gorski, R. A. (1984). Gonadal steroid induction of structural sex differences in the central nervous system. *Annual Review of Neuroscience, 7*(1), 413-442. doi: doi:10.1146/annurev.ne.07.030184.002213

Atance, C. M., Bernstein, D. M., & Meltzoff, A. N. (2010). Thinking about false belief: It's not just what children say, but how long it takes them to say it. *Cognition, 116*(2), 297-301. doi: 10.1016/j.cognition.2010.05.008

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality And Individual Differences, 38*(2), 451-460. doi: 10.1016/j.paid.2004.04.022

Auyeung, B., Ahluwalia, J., Thomson, L., Taylor, K., Hackett, G., O'Donnell, K. J., & Baron-Cohen, S. (2012). Prenatal versus postnatal sex steroid hormone effects on autistic traits in children at 18 to 24 months of age. *Molecular Autism, 3*(1), 17-21. doi: 10.1186/2040-2392-3-17

Auyeung, B., & Baron-Cohen, S. (2012). *Fetal testosterone in mind: Human sex differences and autism.* (F. B. M. de Waal, Ferrari, Pier Francesco Ed. Vol. 1). Cambridge, MA, US: Harvard University Press.

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009). Fetal testosterone and autistic traits. *British Journal of Psychology, 100*(1), 1-22.

Auyeung, B., Baron-Cohen, S., Chapman, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Foetal testosterone and the child systemizing quotient. *European Journal of Endocrinology, 155*(suppl 1), S123-S130. doi: 10.1530/eje.1.02260

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The Autism Spectrum Quotient: Children's Version (AQ-Child). *Journal of Autism and Developmental Disorders, 38*(7), 1230-1240. doi: 10.1007/s10803-007-0504-z

Auyeung, B., Knickmeyer, R., Ashwin, E., Taylor, K., Hackett, G., & Baron-cohen, S. (2012). Effects of fetal testosterone on visuospatial ability. *Archives of Sexual Behavior, 41*(3), 571-581. doi: http://dx.doi.org/10.1007/s10508-011-9864-8

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2009). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of Autism and Developmental Disorders, 39*(11), 1509-1521. doi: 10.1007/s10803-009-0772-x

Bal, E., Harden, E., Lamb, D., Van Hecke, A., Denver, J., & Porges, S. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders, 40*(3), 358-370. doi: 10.1007/s10803-009-0884-3

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254. doi: 10.1037/0033-2909.85.4.845

Baron-Cohen, S. (2010). Chapter 11 - Empathizing, systemizing, and the extreme male brain theory of autism. In S. Ivanka (Ed.), *Progress in Brain Research* (Vol. Volume 186, pp. 167-175): Elsevier.

Baron-Cohen, S., Cassidy, S., Auyeung, B., Allison, C., Achoukhi, M., Robertson, S., . . . Lai, M.-C. (2014). Attenuation of typical sex differences in 800 adults with autism vs. 3,900 controls. *PLoS ONE, 9*(7), e102251. doi: 10.1371/journal.pone.0102251

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition, 21*(1), 37-46. doi: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males? *PLoS Biology, 9*(6), 1-10.

Baron-Cohen, S., Lutchmaya, S., & Knickmeyer, R. (2004). *Prenatal testosterone in mind: Amniotic fluid studies*. Cambridge, MA, US: MIT Press.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29*(5), 407-418. doi: http://dx.doi.org/10.1023/A:1023035012436

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 358*(1430), 361-374. doi: 10.1098/rstb.2002.1206

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-

functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17. doi: 10.1023/a:1005653411471

Berenbaum, S. A., Bryk, K. K., Nowak, N., Quigley, C. A., & Moffat, S. (2009). Fingers as a marker of prenatal androgen exposure. *Endocrinology, 150*(11), 5119-5124. doi: 10.1210/en.2009-0774

Berenbaum, S. A., Bryk, K. L. K., & Beltz, A. M. (2012). Early androgen effects on spatial and mechanical abilities: Evidence from congenital adrenal hyperplasia. *Behavioral Neuroscience, 126*(1), 86-96.

Berenbaum, S. A., & Hines, M. (1992). Early androgens are related to childhood sex-types toy preferences. *Psychological Science, 3*(3), 203-206.

Berenbaum, S. A., & Resnick, S. M. (1997). Early androgen effects on aggression in children and adults with congenital adrenal hyperplasia. *Psychoneuroendocino., 22*(7), 505-515.

Breedlove, S. M. (1994). Sexual differentiation of the human nervous system. *Annual Review of Psychology, 45*, 389-418. doi: 10.1146/annurev.ps.45.020194.002133

Breedlove, S. M., Cooke, B. M., & Jordan, C. L. (1999). The orthodox view of brain sexual differentiation. *Brain, Behavior and Evolution, 54*(1), 8-14.

Brown, G. R., & Dixson, A. F. (1999). Investigation of the role of postnatal testosterone in the expression of sex differences in behavior in infant rhesus macaques (macaca mulatta). *Hormones and Behavior, 35*, 186-194.

Brown, G. R., Nevison, C. M., Fraser, H. M., & Dixson, A. F. (1999). Manipulation of postnatal testosterone levels affects phallic and clitoral development in infant rhesus monkeys. *International Journal of Andrology, 22*(2), 119-128.

Brown, W. M., Hines, M., Fane, B. A., & Breedlove, S. M. (2002). Masculinized Finger Length Patterns in Human Males and Females with Congenital Adrenal Hyperplasia. *Hormones and Behavior, 42*, 380-386. doi: 10.1006/hbeh.2002.1830

Buchanan, C. M., Eccles, J. S., & Becker, J. B. (1992). Are adolescents the victims of raging hormones? Evidence for activational effects of hormones on moods and behavior at adolescence. *Psychological Bulletin, 111*(1), 62-107.

Butterworth, G., & Morissette, P. (1996). Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology, 14*(3), 219-219.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *The American Journal of Psychiatry, 162*(6), 1133-1141.

Chapman, E., Baron-Cohen, S., Auyeung, B., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the empathy quotient (EQ) and the "Reading the Mind in the Eyes" test. *Social Neuroscience, 1*(2), 135-148.

Chawarska, K., & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 39*(12), 1663-1672.

Cho, J. I., Carlo, W. A., Su, X., & McCormick, K. L. (2012). Associations between salivary testosterone and cortisol levels and neonatal health and growth outcomes. *Early Human Development, 88*(10), 789-795. doi: 10.1016/j.earlhumdev.2012.05.002

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101.

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior and Development, 23*(1), 113-118. doi: 10.1016/s0163-6383(00)00032-1

Cook, C. M., & Saucier, D. M. (2010). Mental rotation, targeting ability and Baron-Cohen's empathizing-systemizing theory of sex differences. *Personality And Individual Differences, 49*, 712-716. doi: 10.1016/j.paid.2010.06.010

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science, 24*(5), 723-732. doi: http://dx.doi.org/10.1177/0956797612463582

Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences, 31*, 241-320.

de Cock, M., Maas, Y. G. H., & van de Bor, M. (2012). Does perinatal exposure to endocrine disruptors induce autism spectrum and attention deficit hyperactivity disorders? Review. *Acta Paediatrica, 101*(8), 811-818. doi: 10.1111/j.1651-2227.2012.02693.x

Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH child emotional faces picture set (NIMH-ChEFS): A new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research, 20*(3), 145-156. doi: 10.1002/mpr.343

Engel, S. M., Zhu, C., Berkowitz, G. S., Calafat, A. M., Silva, M. J., Miodovnik, A., & Wolff, M. S. (2009). Prenatal phthalate exposure and performance on the Neonatal Behavioral Assessment Scale in a multiethnic birth cohort. *Neurotoxicology, 30*(4), 522-528. doi: 10.1016/j.neuro.2009.04.001

Eriksson, M. A., Westerlund, J., Anderlid, B. M., Gillberg, C., & Fernell, E. (2012). First-degree relatives of young children with autism spectrum disorders: Some gender aspects. *Research in Developmental Disabilities, 33*(5), 1642-1648. doi: 10.1016/j.ridd.2012.03.025

Evardone, M., & Alexander, G. M. (2009). Anxiety, sex-linked behaviors, and digit ratios (2D:4D). *Archives of Sexual Behavior, 38*(3), 442-455.

Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M. a., & Kumar, V. (2010). Parental demographics and preschool children's theory of mind. *Journal Of Human Ecology, 29*(2), 121-128.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-173.

Fink, B., Manning, J. T., Williams, J. H. G., & Podmore-Nappin, C. (2007). The 2nd to 4th digit ratio and developmental psychopathology in school-aged children.

Personality And Individual Differences, 42(2), 369-379. doi: 10.1016/j.paid.2006.07.018

Forest, M. G., Sizonenko, F. C., Cathiard, A. M., & Bertrand, J. (1974). Hypophyso-gonadal function in humans during the first year of life I. Evidence for testicular activity in early infancy. *Journal of Clinical Investigation, 53*(3), 819-828.

Frick, A., Möhring, W., & Newcombe, N. S. (in press). Development of mental transformation abilities. *Trends in Cognitive Sciences*(0). doi: http://dx.doi.org/10.1016/j.tics.2014.05.011

Friederici, A. D., Pannekamp, A., Partsch, C.-J., Ulmen, U., Oehler, K., Schmutzler, R., & Hesse, V. (2008). Sex hormone testosterone affects language organization in the infant brain. *Neuroreport, 19*(3), 283-286.

Geier, D. A., & Geier, M. R. (2006a). A clinical and laboratory evaluation of methionine cycle-transsulfuration and androgen pathway markers in children with autistic disorders. *Hormone Research, 66*(4), 182-188. doi: http://dx.doi.org/10.1159/000094467

Geier, D. A., & Geier, M. R. (2006b). A clinical trial of combined anti-androgen and anti-heavy metal therapy in autistic disorders. *Neuro endocrinology letters, 27*(6), 833-838.

Geier, D. A., Kern, J. K., King, P. G., Sykes, L. K., & Geier, M. R. (2012). An evaluation of the role and treatment of elevated male hormones in autism spectrum disorders. *Acta Neurobiologiae Experimentalis, 72*(1), 1-17.

Gillberg, C., Cederlund, M., Lamberg, K., & Zeijlon, L. (2006). Brief report: "The autism epidemic". The registered prevalence of autism in a Swedish urban area. *Journal of Autism and Developmental Disorders, 36*(3), 429-435. doi: 10.1007/s10803-006-0081-6

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 38*(8), 1534-1541.

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 36*(2), 169-183.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and Its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*(1), 26-37.

Grimshaw, G. M., Sitarenios, G., & Finegan, J. A. K. (1995). Mental rotation at 7 years: Relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition, 29*(1), 85-100. doi: 10.1006/brcg.1995.1269

Grove, R., Baillie, A., Allison, C., Baron-Cohen, S., & Hoekstra, R. A. (2013). Empathizing, systemizing, and autistic traits: Latent structure in individuals with autism, their parents, and general population controls. *Journal of Abnormal Psychology, 122*(2), 600-609. doi: http://dx.doi.org/10.1037/a0031919

Hahn, N., Jansen, P., & Heil, M. (2010). Preschoolers' mental rotation: Sex differences in hemispheric asymmetry. *Journal of Cognitive Neuroscience, 22*(6), 1244-1250.

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*(3), 843-855. doi: http://dx.doi.org/10.2307/1131954

Hartung, C. M., & Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the DSM-IV. *Psychological Bulletin, 123*(3), 260.

Heil, M., Kavsek, M., Rolke, B., Beste, C., & Jansen, P. (2011). Mental rotation in female fraternal twins: Evidence for intra-uterine hormone transfer? *Biological Psychology, 86*, 90-93.

Hines, M. (2002). Sexual differentiation of the human brain and behavior. In D. W. Pfaff, A. P. Arnold, A. M. Arnold, A. M. Etgen, S. E. Fahrback & R. T. Ruben (Eds.), *Hormones, Brain and Behavior*. San Diego.

Hines, M. (2004). *Brain gender*. New York, NY, US: Oxford University Press.

Hines, M., & Kaufman, F. R. (1994). Androgen and the development of human sex-typical behavior: Rough-and-tumble play and sex of preferred playmates in children with congenital adrenal hyperplasia (CAH). *Child Development, 65*, 1042-1053.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine tuning some resistant rules for outlier labeling. *Journal of American Statistical Association, 82*, 1147-1149.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of American Statistical Association, 81*, 991-999.

Hoekstra, R. A., Bartels, M., Verweij, C. J. H., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatrics & Adolescent Medicine, 161*(4), 372-377.

Hönekopp, J. (2012). Digit ratio 2D:4D in relation to autism spectrum disorders, empathizing, and systemizing: A quantitative review. *Autism Research, 5*(4), 221-230. doi: 10.1002/aur.1230

Hönekopp, J., & Watson, S. S. (2010). Meta-analysis of digit ratio 2D:4D shows greater sex difference in the right hand. *American Journal of Human Biology, 22*(5), 619-630.

Hoyek, N., Collet, C., Fargier, P., & Guillot, A. (2012). The use of the Vandenberg and Kuse Mental Rotation Test in children. *Journal of Individual Differences, 33*(1), 62.

Inozemtseva, O., Matute, E., & Juárez, J. (2008). Learning disabilities spectrum and sexual dimorphic abilities in girls with congenital adrenal hyperplasia. *Journal of Child Neurology, 23*(8), 862-869. doi: 10.1177/0883073808315618

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature, 504*(7480), 427-431. doi: http://dx.doi.org/10.1038/nature12715

Jurewicz, J., & Hanke, W. (2011). Exposure to phthalates: Reproductive outcome and children health. A review of epidemiological studies. *International Journal of Occupational Medicine and Environmental Health, 24*(2), 115-141.

Karabekiroglu, K., Rodopman-Arman, A., Ay, P., Ozkesen, M., Akbas, S., Tasdemir, G. N., . . . Peksen, Y. (2009). The reliability and validity of the Turkish version of the brief infant–toddler social emotional assessment (BITSEA). *Infant Behavior and Development, 32*(3), 291-297. doi: 10.1016/j.infbeh.2009.03.003

Kirchner, J., Hatri, A., Heekeren, H., & Dziobek, I. (2011). Autistic symptomatology, face processing abilities, and eye fixation patterns. *Journal of Autism and Developmental Disorders, 41*(2), 158-167. doi: 10.1007/s10803-010-1032-9

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature, 459*(7244), 257-269.

Knickmeyer, R., & Baron-Cohen, S. (2006). Fetal testosterone and sex differences in typical social development and in autism. *Journal of Child Neurology, 21*, 825-845.

Knickmeyer, R., Baron-Cohen, S., Fane, B. A., Wheelwright, S., Mathews, G. A., Conway, G. S., . . . Hines, M. (2006). Androgens and autistic traits: A study of individuals with congenital adrenal hyperplasia. *Hormones and Behavior, 50*(1), 148-153. doi: 10.1016/j.yhbeh.2006.02.006

Krajmer, P., Špajdel, M., Celec, P., & Ostatníková, D. (2011). Relationship between salivary testosterone levels and empathizing/systemizing in Slovak boys with Asperger syndrome. *Studia Psychologica, 53*(3), 293.

Krajmer, P., Spajdel, M., Kubranska, A., & Ostatnikova, D. (2011). 2D:4D finger ratio in Slovak autism spectrum population. *Bratislavske Lekarske Listy, 112*(7), 377-379.

Lamminmäki, A., Hines, M., Kuiri-Hänninen, T., Kilpeläinen, L., Dunkel, L., & Sankilampi, U. (2012). Testosterone measured in infancy predicts subsequent sex-typed behavior in boys and in girls. *Hormones and Behavior, 61*(4), 611-616. doi: 10.1016/j.yhbeh.2012.02.013

Leeb, R. T., & Rejskind, G. (2004). Here's looking at you, kid! A longitudinal study of perceived gender differences in mutual gaze behavior in young infants. *Sex Roles, 50*(1-2), 1-14.

Lewis, V. (2003). Play and language in children with autism. *Autism, 7*(4), 391.

Lind, S. E., & Bowler, D. M. (2009). Language and theory of mind in autism spectrum disorder: The relationship between complement syntax and false belief task performance. *Journal of Autism and Developmental Disorders, 39*(6), 929.

Lutchmaya, S., & Baron-Cohen, S. (2002). Human sex differences in social and non-social looking preferences, at 12 months of age. *Infant Behavior and Development, 25*(3), 319-325.

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002a). Foetal testosterone and eye contact in 12-month-old human infants. *Infant Behavavior and Development, 25*(3), 327-335.

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002b). Foetal testosterone and vocabulary size in 18- and 24-month-old infants. *Infant Behavavior and Development, 24*(4), 418-424.

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development, 77*(1-2), 23-28.

Luxen, M. F., & Buunk, B. P. (2005). Second-to-fourth digit ratio related to verbal and numerical intelligence and the big five. *Personality And Individual Differences, 39*, 959-966. doi: 10.1016/j.paid.2005.03.016

Mack, C. M., McGivern, R. F., Hyde, L. A., & Denenberg, V. H. (1996). Absence of postnatal testosterone fails to demasculinize the male rat's corpus callosum. *Developmental Brain Research, 95*(2), 252-255. doi: 10.1016/0165-3806(96)00093-4

Macleod, D. J., Sharpe, R. M., Welsh, M., Fisken, M., Scott, H. M., Hutchison, G. R., . . . van den Driesche, S. (2010). Androgen action in the masculinization programming window and development of male reproductive organs. *International Journal of Andrology, 33*(2), 279-287.

MacLusky, N. J., & Naftolin, F. (1981). Sexual differentiation of the central nervous system. *Science, 211*(4488), 1294-1303.

Main, K. M., Schmidt, I. M., & Skakkebæk, N. E. (2000). A possible role for reproductive hormones in newborn boys: Progressive hypogonadism without the postnatal testosterone peak. *Journal of Clinical Endocrinology and Metabolism, 85*(12), 4905-4907. doi: 10.1210/jc.85.12.4905

Manning, J. T., Baron-Cohen, S., Wheelwright, S., & Fink, B. (2010). Is digit ratio (2D:4D) related to systemizing and empathizing? Evidence from direct finger measurements reported in the BBC internet survey. *Personality And Individual Differences, 48*(6), 767-771. doi: http://dx.doi.org/10.1016/j.paid.2010.01.030

Manning, J. T., Bundred, P. E., Newton, D. J., & Flanagan, B. F. (2003). The second to fourth digit ratio and variation in the androgen receptor gene. *Evolution and Human Behavior, 24*, 399-405.

Markram, H., Rinaldi, T., & Markram, K. (2007). The intense world syndrome--an alternative hypothesis for autism. *Frontiers in neuroscience, 1*(1), 77-96. doi: http://dx.doi.org/10.3389/neuro.01.1.1.006.2007

Marmor, G. S. (1975). Development of kinetic images: When does the child first represent movement in mental images? *Cognitive Psychology, 7*(4), 548-559. doi: 10.1016/0010-0285(75)90022-5

Marmor, G. S. (1977). Mental rotation and number conservation: Are they related? *Developmental Psychology, 13*(4), 320.

Martin, N. A., & Brownell, R. (2010). Expressive One-Word Picture Vocabulary Test-Fourth Edition. Novato, CA: Academic Therapy Publications.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2003). Mental health of children and adolescents in Great Britain. *International Review of Psychiatry, 15*(1/2), 185.

Milne, E., White, S., Campbell, R., Swettenham, J., Hansen, P., & Ramus, F. (2006). Motion and form coherence detection in autistic spectrum disorder: Relationship to motor control and 2:4 digit ratio. *Journal of Autism and Developmental Disorders, 36*(2), 225-237. doi: 10.1007/s10803-005-0052-3

Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science, 19*(11), 1063-1066.

Nevison, C. M., Brown, G. R., & Dixson, A. F. (1997). Effects of altering testosterone in early infancy on social behaviour in captive yearling rhesus monkeys. *Physiology and Behavior, 62*(6), 1397-1403.

O'Neill, D. K., & Happe, F. G. E. (2000). Noticing and commenting on what's new: Differences and similarities among 22-month-old typically developing children, children with Down syndrome and children with autism. *Developmental Science, 3*(4), 457-478.

Okten, A., Kalyoncu, M., & Yaris, N. (2002). The ratio of second-and fourth-digit lengths and congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Early Human Development, 70*(1-2), 47-54.

Pang, S., Levine, L. S., Cederqvist, L. L., Fuentes, M., Riccardi V, M., Holcombe, J. H., . . . Merkatz, I. (1980). Amniotic fluid concentrations of delta-5 steroids and delta-4 steroids in fetuses with congenital adrenal hyperplasia due to 21

hydroxylase deficiency and anencephalic fetuses. *Journal of Clinical Endocrinology and Metabolism, 51*(2), 223.

Pasterski, V., Geffner, M. E., Brain, C., Hindmarsh, P., Brook, C., & Hines, M. (2005). Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Development, 76*, 264-278.

Pasterski, V., Geffner, M. E., Brain, C., Hindmarsh, P., Brook, C., & Hines, M. (2011). Prenatal hormones and childhood sex segregation: Playmate and play style preferences in girls with congenital adrenal hyperplasia. *Hormones and Behavior, 59*(4), 549-555. doi: 10.1016/j.yhbeh.2011.02.007

Pasterski, V., Mindmarsh, P., Geffner, M., Brook, C., Brain, C., & Hines, M. (2007). Increased aggression and activity level in 3- to 11-year-old girls with congenital adrenal hyperplasia (CAH). *Hormones and Behavior, 52*, 368-374.

Peper, J. S., & Koolschijn, C. M. P. (2012). Sex steroids and the organization of the human brain. *The Journal of Neuroscience, 32*(20), 6745.

Perner, J., Kloo, D., & Rohwer, M. (2010). Retro- and prospection for mental time travel: Emergence of episodic remembering and mental rotation in 5- to 8-year old children. *Consciousness and Cognition, 19*(3), 802-815. doi: 10.1016/j.concog.2010.06.022

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125-137. doi: 10.1111/j.2044-835X.1987.tb01048.x

Peters, M., Manning, J. T., & Reimers, S. (2007). The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Archives of Sexual Behavior, 36*(2), 251-260.

Phoenix, C. H., Goy, R. W., Gerral, A. A., & Young, W. C. (1959). Organizing action of prenatally administered testosterone propionate on the tissue mediating mating behavior in the female guinea pig. *Endocrinology, 65*(163-196), 163.

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry, 68*(1), 101-109. doi: 10.1001/archgenpsychiatry.2010.113

Puts, D. A., McDaniel, M. A., Jordan, C. L., & Breedlove, S. M. (2008). Spatial ability and prenatal androgens: Meta-analyses of congenital adrenal hyperplasia and digit ratio (2D:4D) studies. *Archives of Sexual Behavior, 37*(1), 100-111.

Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science, 19*(11), 1067-1070.

Quinton, S. J., Smith, A. R., & Joiner, T. (2011). The 2nd to 4th digit ratio (2D:4D) and eating disorder diagnosis in women. *Personality And Individual Differences, 51*(4), 402-405. doi: 10.1016/j.paid.2010.07.024

Rice, K., Moriuchi, J. M., Jones, W., & Klin, A. (2012). Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(3), 238.

Ruta, L., Ingudomnukul, E., Taylor, K., Chakrabarti, B., & Baron-Cohen, S. (2011). Increased serum androstenedione in adults with autism spectrum conditions. *Psychoneuroendocrinology, 36*(8), 1154-1163. doi: http://dx.doi.org/10.1016/j.psyneuen.2011.02.007

Saenz, J., & Alexander, G. M. (2013a). Digit ratios (2D:4D), postnatal testosterone and eye contact in toddlers. *Biological Psychology, 94*(1), 106-108.

Saenz, J., & Alexander, G. M. (2013b). Postnatal testosterone levels and disorder relevant behavior in the second year of life. *Biological Psychology, 94*(1), 152-159.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701-703. doi: 10.1126/science.171.3972.701

Smith, A. (2009). The empathy imbalance hypothesis of autism: A theoretical approach to cognitive and emotional empathy in autistic development. *The Psychological Record, 59*(3), 489-510.

Soper, D. S. (2014). A-priori sample size calculator for multiple regression [software]. from http://www.danielsoper.com/statcalc

Stoner, S. B., & Spencer, W. B. (1983). Sex differences in expressive vocabulary of head start children. *Perceptual and Motor Skills, 56*(3), 1008-1008. doi: 10.2466/pms.1983.56.3.1008

Swan, S. H., Liu, F., Hines, M., Kruse, R. L., Wang, C., Redmon, J. B., . . . Weiss, B. (2010). Prenatal phthalate exposure and reduced masculine play in boys. *International Journal of Andrology, 33*(2), 259-269.

Swan, S. H., Main, K. M., Liu, F., Stewart, S. L., Kruse, R. L., Calafat, A. M., . . . the Study for Future Families Research, T. (2005). Decrease in Anogenital Distance among Male Infants with Prenatal Phthalate Exposure. *Environmental Health Perspectives, 113*(8).

Tapanainen, J., Koivisto, M., Vihko, R., & Huhtaniemi, I. (1981). Enhanced activity of the pituitary-gonadal axis in premature human infants. *The Journal of clinical endocrinology and metabolism, 52*(2), 235-238.

van Anders, S. M., Vernon, P. A., & Wilbur, C. J. (2006). Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Hormones and Behavior, 49*(3), 315-319. doi: 10.1016/j.yhbeh.2005.08.003

van den Driesche, S., Scott, H. M., MacLeod, D. J., Fisken, M., Walker, M., & Sharpe, R. M. (2011). Relative importance of prenatal and postnatal androgen action in determining growth of the penis and anogenital distance in the rat before, during and after puberty. *International Journal of Andrology, 34*, e578-e586. doi: 10.1111/j.1365-2605.2011.01175.x

Venkatasubramanian, G., Arasappa, R., Rao, N. P., & Gangadhar, B. N. (2011). Digit ratio (2D:4D) asymmetry and schneiderian first rank symptoms: Implications for cerebral lateralisation theories of schizophrenia. *Laterality, 16*(4), 499.

Voyer, D., & Hou, J. (2006). Type of items and the magnitude of gender differences on the mental rotations test. *Canadian Journal of Experimental Psychology, 60*(2), 91-100.

Vuoksimaa, E., Jaakko, K., Kremen, W. S., Hokkanen, L., Viken, R. J., Tuulio-Henriksson, A., & Rose, R. J. (2010). Having a male co-twin masculinizes mental rotation performance in females. *Psychological Science, 21*(8), 1069-1071.

Wakabayashi, A., & Nakazawa, Y. (2010). On relationships between digit ratio (2D:4D) and two fundamental cognitive drives, empathizing and systemizing, in Japanese sample. *Personality And Individual Differences, 49*(8), 928-931. doi: http://dx.doi.org/10.1016/j.paid.2010.07.032

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *The Journal of Genetic Psychology, 166*(3), 297-312.

Wallen, K., Maestripieri, D., & Mann, D. R. (1995). Effects of neonatal testicular suppression with a GnRH Antagonist on social behavior in group-living juvenile rhesus monkeys. *Hormones and Behavior, 29*(3), 322-337. doi: http://dx.doi.org/10.1006/hbeh.1995.1023

Whitehouse, A. J. O., Mattes, E., Maybery, M. T., Sawyer, M. G., Jacoby, P., Keelan, J. A., & Hickey, M. (2012). Sex‐specific associations between umbilical cord blood testosterone levels and language delay in early childhood. *Journal of Child Psychology and Psychiatry, 53*(7), 726-734. doi: http://dx.doi.org/10.1111/j.1469-7610.2011.02523.x

Whitehouse, A. J. O., Maybery, M. T., Hart, R., Mattes, E., Newnham, J. P., Sloboda, D. M., . . . Hickey, M. (2010). Fetal androgen exposure and pragmatic language

ability of girls in middle childhood: Implications for the extreme male-brain theory of autism. *Psychoneuroendocrinology, 35*, 1259-1264.

Williams, J. H. G., Greenhalgh, K. D., & Manning, J. T. (2003). Second to fourth finger ratio and possible precursors of developmental psychopathology in preschool children. *Early Human Development, 72*(1), 57-65. doi: 10.1016/S0378-3782(03)00012-4

Zahn-Waxler, C. (1993). Warriors and worriers: Gender and psychopathology. *Development and Psychopathology, 5*(1-2), 79.

Zahn-Waxler, C., Shirtcliff, E. A., & Marceau, K. (2008). Disorders of childhood and adolescence: Gender and psychopathology. *Annual Review of Clinical Psychology, 4*(1), 275-303. doi: doi:10.1146/annurev.clinpsy.3.022806.091358

Zuloaga, D. G., McGivern, R. F., & Handa, R. J. (2009). Organizational influence of the postnatal testosterone surge on the circadian rhythm of core body temperature of adult male rats. *Brain Research, 1268*(0), 68-75. doi: 10.1016/j.brainres.2009.02.048