EMERGENT RELATIONSHIPS BETWEEN EMPATHIZING AND SYSTEMIZING SKILLS IN AUTISTIC AND TYPICAL DEVELOPMENT

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jeffrey M. Valla

August 2013

EMERGENT RELATIONSHIPS BETWEEN EMPATHIZING AND SYSTEMIZING SKILLS IN AUTISTIC AND TYPICAL DEVELOPMENT

Jeffrey M. Valla, Ph. D.

Cornell University 2013

Empathizing-systemizing (E-S) theory attempts to explain individual cognitive variation in terms of an extrapolation of the autism spectrum. In this view, autism is an extreme version of the human brain in terms of its "male-ness," the prototypical male brain being organized – via prenatal androgen exposure – for 'systemizing,' or detail-oriented, observation-based, 'lawful, finite, and deterministic' rule-making. The 'extreme female brain,' by contrast, is organized for 'empathizing,' the ability to accurately attribute intentionality to others. The full range of these attributes constitutes the E-S spectrum, within which the general population is normally distributed. E-S theory helps explain individual and sex differences with real world importance, such as the social idiosyncrasies characterizing many individuals in systemizing fields (e.g., engineering, computer science) and the underrepresentation of women in STEM. Implicit in E-S theory is the assumption that empathizing and systemizing domains co-vary; that social tradeoffs come with a detail-oriented cognitive "style". This assumption is recognizable in the longstanding search for singular causes of autism that will explain social and non-social symptoms; and in everyday presumptions about the social skills of individuals excelling in numerical and spatial domains. We begin with an overview of the E-S covariance literature and the three works comprising this dissertation, followed by individual chapters discussing each work in depth. The first is an empirical investigation indicating that E-S co-variance assumptions apply more to males than females, in whom the two domains are orthogonal. These sex-

dependent patterns extend to college major, suggesting an ability breadth alternative to the absolute ability account of female STEM underrepresentation. The second reviews cognitive sex differences studies utilizing second-to-fourth digit length ration (2D/4D) as a biomarker of prenatal testosterone exposure, and argues that the biological linchpin of E-S theory – the brain organization effects of prenatal testosterone exposure – may also be sex-dependent. The third synthesizes the literature on empathizing-systemizing co-variance across clinical, sub-clinical, and typically developing populations into a model of developmentally emergent E-S co-variance. In each case co-variance assumptions are confirmed and challenged, providing novel insights about typical and autistic development. In the concluding chapter this model is used to account for the findings of the first two works, and future empirical directions are suggested based on the collective insights of all three works.

BIOGRAPHICAL SKETCH

Jeffrey Valla received a B.S. in Human Development in 2004, a M.S. in Developmental Psychology in 2007, and a Ph.D. in Developmental Psychology in 2013, all from the College of Human Ecology, Cornell University, Ithaca, NY.

For Justin

ACKNOWLEDGMENTS

Stephen Ceci, Wendy Williams, Matthew Belmonte, Steve Robertson, Andrew Barsky, Victoria Dahl, Keith Yoder, Barbara Ganzel, Grace Chen, Jeffrey Maendel, Barbara Koslowski, Mom, Dad, Jenny, Justin, Andy, Benjamin, Kayla, Killian, and Katie.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

THE EMPATHIZING-SYSTEMIZING CO-VARIANCE HYPOTHESIS

1.1 Introduction

The topical focus of this thesis is the hypothesis, hereafter referred to as the "Empathizing-Systemizing (E-S) Co-variance Hypothesis", that individual differences in so-called "systemizing" domains of ability – those favoring a detail-oriented cognitive "style" (e.g., spatial, numerical, and technical domains) – are inversely related to individual differences in the ability to make accurate social inferences about others' intentions ("empathizing", "Theory of Mind", or "mindreading"). As the reader will see, the E-S Co-variance Hypothesis manifests in many forms in the empirical literature, from underlying causes of autism to neurobiological accounts of the underrepresentation of women in science, but for introductory purposes readers may find it helpful to conceptualize the E-S Co-variance Hypothesis in terms of a common lay psychological theory: that prowess in fields like mathematics and engineering comes hand-in-hand with social ineptitude.

The primary goal of the present thesis is to assess the veracity of the E-S Co-variance Hypothesis, through the presentation and cross-referenced discussion of three original, published works. At face, these three works appear unrelated, varying widely in specific focus, approach, and level of analysis; but each holds direct relevance to E-S Co-variance Hypothesis, and together they provide a cross-section of its many levels of prediction. Thus, in presenting and discussing these three tangentially related works, the reader will be provided with the multiple theoretical perspectives, methodological approaches, sample populations, and levels of analysis needed to fully evaluate the veracity of the E-S Co-variance Hypothesis.

1.2 Work One: Sex and Field Differences in Empathizing-Systemizing Co-variance

In the realm of normative cognitive variation, the E-S Co-variance Hypothesis takes the form of cognitive sex difference predictions, namely that males are more likely than females to exhibit the high systemizing/low empathizing extremity of the inverse E-S Co-variance relationship. This prediction has, in turn, been extrapolated into the realm of sex differences in field of study/occupational choices, and an account of female underrepresentation in STEM disciplines.

Addressing this version of the E-S Co-variance Hypothesis is the first work presented below: an empirical investigation of sex and field of study differences in Empathizing, Systemizing, and Empathizing-Systemizing Co-variance, with a normative adult sample (Valla et al., 2010). In this correlational study, undergraduate students from a wide range of academic disciplines completed a variety of systemizing- and empathizing-related cognitive, social cognitive, and self-report personality measures. The main finding was that Empathizing-Systemizing Co-variance is a *sex-dependent* phenomenon, more pronounced in males than in females. Moreover, these sex-dependent patterns extended to field of study; males in systemizing disciplines were equally notable for their social cognitive difficulties and detail-oriented cognitive 'style', whereas the detail-oriented cognitive style of females in these disciplines did not appear to come hand-in-hand with social cognitive difficulties.

1.3 Work Two: Brain Organization Theory, Digit Ratio (2D/4D), and Sex Differences in Preferences and Cognition

The underlying basis of the sex differences explored in the first work is thought to be hemispheric lateralization, induced by prenatal androgen exposure. Addressing this neurobiological level of the E-S Co-variance Hypothesis is the second work presented below, a

review and synthesis of one particular subset of the literature on prenatal hormonal influences on cognitive sex differences in systemizing and empathizing domains (Valla & Ceci, 2011): studies using 2nd-to-4th digit length ratio (2D:4D) as a proxy measure of prenatal androgen exposure. In this second work it is first argued that 2D:4D evidence generally supports the idea of cognitive sex differences, and prenatal androgen exposure as the biological impetus for these sex differences. It is then argued that the sex-dependent patterns of E-S Co-variance seen in the first work – greater E-S Co-variance in males than in females – may be the result of sex-dependent developmental trajectories that begin with prenatal hormone lateralization, resulting in downstream effects on preferences for and ability in empathizing and systemizing domains, starting in early development and increasing with age. This second work therefore extends support for E-S Co-variance to prenatal biological factors and early development, laying the groundwork for the developmental model of E-S Co-variance introduced in the third work.

1.4 Work Three: Co-variation of Social and Non-social 'Autistic' Traits, in Autistic and Typical Development: A Review and Developmental Synthesis

The third work ties the concept of E-S Co-variance to the realm of autism and its sub-clinical manifestations. Specifically, in the context of autism the E-S Co-variance Hypothesis manifests in the idea that autistic social difficulties and hyper-focused, systemizing-centric cognitive 'style' are related, such that individuals on the spectrum represent an extreme version of the high systemizing/low empathizing "male" pattern argued in the first two works. This hypothesis of autistic social/non-social trait co-variance, sometimes referred to as the "Kanner Hypothesis" in the autism literature, is in fact a longstanding debate in autism research, as the question of social/non-social trait co-variance ultimately comes down to whether there is a single underlying cause of autism.

The third work addresses this clinical version of the E-S Co-variance Hypothesis, and ties it to the realms of normative cognitive variation and its development discussed in the first two works. More specifically, this third work begins with a review of the evidence regarding the factor structure of autistic social and non-social traits, including those falling under the categories of empathizing and systemizing, in both clinical and typically developing populations. The gist of this review is that there is overall support for the E-S Co-variance Hypothesis, as well as some indication that E-S Co-variance 1) increases with extremity in empathizing and systemizing domains, such that males are more likely to exhibit co-variance than females, and autistic individuals are more likely to exhibit co-variance than typically developing individuals; and 2) increases with development, such that co-variance increases with age, especially in autistic and typically developing male populations. Based on this review, a developmental model is then proposed to account for these patterns of E-S co-variance in behavior, and tie them to neurobiological development.

CHAPTER 2

SEX AND FIELD DIFFERENCES IN EMPATHIZING-SYSTEMIZING CO-VARIANCE

2.1 Introduction

Baron-Cohen's "Extreme Male Brain" (EMB) and "Empathizing-Systemizing" (E/S) theories attempt to explain an entire range of individual cognitive and social variation, clinical and normative, in terms of an extrapolation of the autism spectrum. Characteristics said to vary along this continuum include deficits in understanding intentionality, or Theory of Mind (ToM); a bias toward local over global information processing, or weak central coherence (WCC); deficits in future-oriented cognitive flexibility; and restricted interests and repetitive behaviors (Baron-Cohen, 1995; Frith & Happé, 1994; Ozonoff, Pennington, & Rogers, 1991). In this view, autism (along with Asperger syndrome, Pervasive Developmental Disorder Not Otherwise Specified, and the sub-clinical idiosyncrasies of many autism family members, or the Broader Autism Phenotype; Baron-Cohen & Hammer, 1997a; Dawson et al., 2002; Happé et al., 2001; Lord, Cook, Leventhal, & Amaral, 2000; Piven, 1999) is an extreme version of the human brain in terms of its "male-ness," the prototypical male brain being organised for superior "systemizing" (Baron-Cohen, 2002; Baron-Cohen & Hammer, 1997b), or processes of observation-based rule-making that reduce the world to a series of "lawful, finite, and deterministic" rules. The "extreme female brain", by contrast, is organized for "empathizing", the ability to accurately attribute and affectively respond to others' mental states (Baron-Cohen, 2002).

The full range of these attributes, from high to low empathizing and from high to low systemizing, constitutes the entire E/S spectrum, within which the general population is said to be distributed normally (Baron-Cohen & Hammer, 1997b). One's position on this continuum is

thought to be determined partly by levels of testosterone exposure during gestation (Baron-Cohen, Lutchmaya, & Knickmeyer, 2004). Most recently, Chakrabarti et al. (2009) have found support for this connection, linking sex steroid genes to autistic traits, and to empathizing, as measured by the Autism Spectrum Quotient (AQ) and Empathizing Quotient, respectively; the former is a self-report survey used for measuring sub-clinical levels of autistic traits, the latter a self-report survey of empathic tendencies.

Still unclear is whether empathizing and systemizing are independent or dependent abilities or, in other words, whether the E/S factor space is single- or multi-axis in nature (Carroll & Chiew, 2006). More than a trivial distinction, if the E/S factor space is a single axis along which all individuals are distributed, as it was initially described (Baron-Cohen & Hammer, 1997b), then systemizing ability should carry with it *tradeoffs* in empathizing, and vice versa.

An alternative model, explicitly (and somewhat paradoxically) depicted by Baron-Cohen in his EMB synthesis and subsequent work (Baron-Cohen, 2002), describes the E/S spectrum as a two-factor space where empathizing and systemizing vary *independently*, an extreme male brain arising from concurrently low empathizing and high systemizing abilities, an extreme female brain arising from the opposite pattern. The major implication is that there would be no tradeoffs between systemizing and empathizing. If empathizing and systemizing are independent, reason Carroll and Chiew, then so too are ToM (an empathizing ability) and WCC (a systemizing ability). If such were the case, then E/S theory would have to be reframed, as the posited covariance of empathizing and systemizing is what has allowed E/S theory to explain social and non-social cognitive variations simultaneously. A lack of covariance would also be a blow to theories attempting to account for autism spectrum conditions (ASC) with a single-cause explanation, and the broader idea of ASC's arising from coincidence of independent factors, both

distal (genetics), and proximal (cognitive manifestations) (Happé, Ronald, & Plomin, 2006), might begin to seem more likely.

2.1.1 Empathizing and Systemizing in the Normative Population

Exploring cognitive covariance in the general population allows us to test the predictions of E/S theory and, indirectly, to test whether there may be a central factor underlying autism. In a sense, testing the general population makes a more sound strategy than testing the ASC population would do: exploring covariance within ASC would, by clinical definition, mean testing a population in which the traits in question co-exist within each individual; thus if there is a disjunction between empathizing and systemizing, assessing the general population offers the best chance of discovering it (Happé et al., 2006). To date, evidence of covariance and sex differences from normative populations has been mixed.

Jarrold et al. (2000) found a negative correlation between ToM (tested via the "Reading the Mind in the Eyes" test, or RMET; Baron-Cohen, Wheelwright, Raste, Hill, & Plumb, 2001a) and WCC (tested via the Embedded Figures Test, or EFT) overall, as predicted by a single-axis E/S spectrum. In their sample of 30 individuals of each sex, the relationship reached statistical significance in women but not in men, suggesting a partially sex-dependent ToM-WCC covariance pattern. None of the EMB model's predicted sex differences in RMET and EFT scores were found. This combination of results supported the single-axis E/S model, but not EMB theory.

Carroll and Chiew (2006), in contrast, found no relationship between the RMET and either of two WCC measures (the EFT and a Block Design test), between systemizing ability and the RMET, or between the EQ and Systemizing Quotient (SQ), supporting a dual-axis E/S spectrum for both sexes. Also in contrast with Jarrold et al., they found sex differences on all

measures except for the EFT. Similarly supporting empathizing and systemizing independence and predicted sex differences, Walter, Dassonville, and Bochler (2008) found that resistance to visual contextual illusions (an aspect of WCC) correlated with SQ scores, while SQ and EQ scores were independent of each other but related to sex in the predicted direction.

Hoekstra, Bartels, Cath, and Boomsma (2008) factor analyzed the five subcategories of autistic symptoms comprising the AQ (Attention Switching, Communication, Imagination, Social Skills, detail/pattern orientation), and identified two underlying factors, broadly termed Social Interaction (encompassing the first four subcategories) and Details/Patterns, using both normative and clinical populations. Although this result supports the dual-axis E/S model, the questions from the Social Interaction subcategories do overlap with each other much more than with the details/patterns questions; attention switching, for instance, is partly posed in the context of keeping track of multiple conversations simultaneously, whereas orientation to detail is never posed in a social context. The apparent two-factor symptom structure thus may be an artifact of the AQ's having been designed with an empathizing- systemizing dichotomy implicitly in mind.

Voracek and Dressler (2006) found no correlations between the SQ, EQ, and RMET in females, whereas males manifested correlations opposite to those predicted, positive between the RMET and SQ and between the EQ and SQ, supporting a sex-dependent E/S factor space and suggesting that males may solve empathizing problems by applying systemizing skills—a point to which we shall return. Thus, evidence exists for three versions of the E/S factor space: (1) dependence between empathizing and systemizing, (2) independence between these abilities, and (3) both of these covariance patterns, the former in males, the latter in females.

2.1.2 Empathizing, Systemizing, and Field of Study

Separate from sex difference and covariance predictions are predicted field-of-study differences, specifically the hypothesis that populations in mathematically intensive fields (mathematics, engineering, computer science, physical sciences) will exhibit ‘‘extreme male brain’’ characteristics, while populations in biological sciences

and medicine will exhibit ‘‘male brain’’ characteristics to a lesser extent than the more mathematically intensive fields, but more so than social sciences and humanities.

Carroll and Chiew (2006) found predicted field differences on the RMET, Block Design test, and SQ, but no differences on the AQ or EQ, partially supporting predictions of E/S theory. Contrary to expectations, a sex-by-field interaction arose for the EQ wherein male humanities students scored significantly lower than male science students, female science students, and female humanities students, the latter three being no different from one another. Austin (2005), in contrast, did find the predicted difference in AQ scores between students in mathematical and physical sciences, and those in all other fields, while Baron-Cohen, Wheelwright, Skinner, Martin, and Clubley, (2001b) found that among systemizing fields, mathematicians scored higher on the AQ than engineers, physicists, and computer scientists (who did not differ), and that individuals in these four fields scored significantly higher than individuals in medical and biological sciences (who did not differ). Moreover, individuals from these six fields scored higher than individuals from social sciences and humanities. Thus, degrees of E/S difference may exist even between fields traditionally grouped under the systemizing label.

2.1.3 Digit Ratio

An increasingly popular method for measuring prenatal testosterone exposure, if indirectly, is measuring individuals’ second-to-fourth digit length ratio (2D:4D). Since the

former digit's length is correlated with prenatal oestrogen level, and the latter with prenatal testosterone level, the 2D:4D ratio hypothetically indicates relative levels of prenatal testosterone and oestrogen exposures in utero (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004), a lower 2D:4D being more masculine. Manning, Baron-Cohen, Wheelwright, and Sanders (2001) found support for EMB in a digit ratio study in which individuals with ASC, and their immediate family members, had lower 2D:4D than normal controls. The effect of prenatal testosterone on systemizing skills may also be sex dependent (Falter, Arroyo, & Davis, 2006; Kempel et al., 2005). Evidence indicates a quadratic, inverted-U relationship between 2D:4D and systemizing in males, with the optimal 2D:4D for systemizing being in the high (feminine) male range, and an inverse relationship in females (Falter et al., 2006; Kempel et al., 2005; Romano, Leoni, & Saino, 2006), implying that EMB predictions of prenatal testosterone exposure might apply only to females. Voracek and Dressler (2006), however, found no relationships between 2D:4D and the RMET, EQ, or SQ in either sex.

2.2 Current Study

Though the structural nature and veracity of group difference predictions of E/S theory are unclear, what is clear is that questions of covariance, and sex and field differences among ASC-related attributes in the general population remain open. Covariance is most completely assayed by (1) taking data from the *same* participants on a wide array of ASC-relevant skills and attributes (RMET, EFT, 2D:4D, etc.), rather than relying on parallels between studies that only partially overlap methodologically; (2) using both standard and novel measures, so that evidence does not rest solely on the most frequently-used measures (EFT and RMET, e.g.); and (3) testing for differences in empathizing and systemizing between social sciences and humanities, and between humanities and fine arts – comparisons that, thus far incompletely explored, may

challenge the assumption that the high-empathizing end of the E/S factor space (high empathizing with separate but concurrent low systemizing in the dual-axis version; high empathizing with correlated low systemizing in the single-axis version) does not exhibit fine-grained differences akin to those observed between systemizing fields. Intuitively, this fine-grained subdivision seems possible, as social science involves constructing predictable systems to explain behaviors that are initially "observed" through empathic conjectures, whereas humanities and fine arts are more heavily conceptual and emotive than systematic.

These goals and new questions are addressed by assaying covariance among, and sex and field differences in, the RMET, EFT, AQ, and 2D:4D using a single adult participant pool. These frequently used measures are complemented with the Benton Face Recognition Test (Benton, Hamsher, Varney, & Spreen, 1983), face recognition being an area of autistic deficit not yet explored in the context of normal cognitive variation and E/S theory. Two further novel measures are presented in the motivating and ecologically valid context of a video game: The first of these is psychophysical threshold for perception of dot motion coherence, a measure of WCC; performance on this measure is a function of integrative or global processing, rather than a function of lack of global processing as in the cases of the EFT and Block Design test. The second measure is go/no-go motor inhibition, an aspect of executive functioning and ASC deficit not typically examined in relation to the aforementioned measures in normative populations.

On the theory of inversely related empathizing and systemizing, the dual-axis model of E/S factor space notwithstanding, we hypothesized positive correlations between RMET accuracy, EFT response latency, Benton accuracy, go/no-go response speed, and 2D:4D, and negative correlations between these measures and the motion coherence threshold. Assuming replication of Hoekstra et al.'s two-component (Social Interaction, Details) model of AQ

subdomain scores, we expected AQ Social Interaction composite scores to be negatively correlated with RMET accuracy, EFT latency, Benton accuracy, go/no-go response speed, and 2D:4D, and positively correlated with motion coherence threshold; we expected AQ Details composite scores to exhibit the inverse pattern.

We expected women to outperform men on the RMET, Benton, go/no-go, and motion coherence measures, and to have higher 2D:4D than men, and we expected men to outperform women on the EFT, and to have higher AQ scores. Field-of-study differences were hypothesized to follow a spectrum embedded within the E/S factor space, from fine arts and humanities at the high-empathizing-low-systemizing edge of the space to social sciences, to biological and medical sciences, to physical sciences, mathematics, and engineering at the high-systemizing/low-empathizing edge.

2.3 Methods

144 university students (65 men, 79 women; M = 20.2y, SD = 1.86y) were recruited from undergraduate classes. Participants were pre-screened, excluding personal history of neurological, sensory, and/or endocrine disorders, pregnancy, transgender identity, and personal or family history of mood disorder, schizophrenia, attention deficit disorder, or ASC. The study protocol was approved by the Cornell University Institutional Review Board. Participants completed a demographic survey (sex, date of birth, field of study). Fields of study were coded on a 5-point scale with fine arts (visual arts, theatre, dance, film, &c.) as 1, humanities (literature, linguistics, history, religion, philosophy, &c.) as 2, social sciences (economics, sociology, psychology, anthropology, political science, &c.) as 3, biological sciences and medicine as 4, and physical sciences (physics, chemistry, geology, &c.) and other mathematically-intensive fields (computer science, mathematics, systems science, &c.) as 5.

Participants completed the AQ, RMET, a computer-based EFT (Witkin, Oltman, Raskin, & Karp, 1971), the Benton, and tests of motion coherence and go/no-go motor inhibition embedded in a video game. To measure motion coherence, participants viewed a field of 200 randomly distributed white dots ("stars"), against a black background, on a 332mm x 210mm display. During each trial (3 blocks, 140 trials/block), either a certain percentage of "stars" would move coherently (signal) to the left or right (randomly chosen for each trial) at 16°/s while the remaining "stars" (noise) would move in random directions, or no coherent motion would be present. Each dot remained for exactly four 60 Hz video frames. Participants were instructed to press a right or left arrow key corresponding to the direction of perceived motion, and visual feedback was given to indicate performance accuracy. Responses were accepted until 0.5s after the end of the trial. Trials ended when a response was issued, or after 2s plus a random and uniformly distributed interval between 0 and 250ms. Motion coherence threshold was computed via maximum-likelihood parameter estimation by sequential testing (PEST; Pentland, 1980).

For the motor inhibition measure, at random intervals during the game the player was offered the chance to open a "wormhole" through which a friend or a foe, distinguished by shape, would appear. Targets were presented after the participant signaled readiness by pressing a key to open the wormhole. The participant was to "shoot" (via a key press) foes but not friends. Reaction times were measured up to a maximum of 800 ms following the emergence of a foe.

The AQ was scored similarly to Austin's (2005) four-point Likert scale (1-4) rather than Baron-Cohen's 0/1 scoring, except instead of a 1-4 scale a symmetric scale of -2, -1, +1, +2 was used. The five AQ subcategory scores, attention switching (AQAttSw), communication

(AQComm), details/patterns (AQDet), imagination (AQImag), and social skills (AQSS) were also recorded.

Each participant's right hand was scanned at a resolution of 300 dpi for 2D:4D measurement. The horizontal and vertical coordinates of the tip and of the midpoint of the basal crease of each participant's index (2D) and ring (4D) fingers were noted in mm, finger lengths were calculated, and 2D:4D ratios were computed using these finger lengths.

2.4 Results

Means and standard deviations for each measure, by sex and field of study, are summarized in Table 8. Outliers for each measure, defined as being greater than 3 times the interquartile range, were removed from the data; 1 outlier was removed from Benton scores, 1 from digit ratios, and 3 from EFT latencies. AQ scores were unavailable for 6 participants who failed to complete all questions. Because of technological issues during ongoing software development for the video game, motion coherence data were available from only 64 participants, and go/no-go data from only 128 participants. All measures were normally distributed according to Kolmogorov-Smirnov goodness-of-fit tests except for EFT latencies, which were normalized using a logarithmic transformation.

2.4.1 AQ Subdomain Principal Components Analyses

AQ subdomain scores AQAttSw, AQComm, AQDet, AQImag, and AQSS from the entire sample were subjected to a principal components analysis (PCA); all components with eigenvalues greater than one were extracted (Kaiser method) (Table 2). In replication of Hoekstra et al., this analysis yielded two principal components: the first, accounting for 44% of the total variance in AQ subdomain scores, was consistent with Hoekstra et al.'s Social Interaction factor, loading on subdomains AQAtt, AQCom, AQSS, and AQImag; the second,

accounting for 20% of the variance, was consistent with Hoekstra et al.'s Details factor, comprised of the AQDet subdomain alone. Based on this PCA, Social Interaction (AQSI) and Details (AQDet) composite scores were calculated for each participant, to be correlated with behavioral measures.

Table 1. Summary of Means and Standard Deviations, by Sex and Field of Study

Field	Sex		AQ (Full)	AQSI	AQDet	Benton	RMET	2D:4D	EFT RT	Motion Threshold	Go/no-go RT
1	Female	Mean	-23.78	-22.22	-1.56	47.88	27.22	.97	5869	.235	825
		SD	24.25	22.59	5.34	1.81	2.28	.03	1575	.083	186
	Male	Mean	-25.00	-26.75	1.75	48.50	29.75	.95	5221	.359	807
		SD	20.80	14.38	9.29	4.65	2.63	.04	1731	.001	217
	Total	Mean	-24.15	-23.62	-0.54	48.08	28.00	.96	5653	.253	819
		SD	22.37	19.92	6.57	2.84	2.58	.04	1580	.089	187
2	Female	Mean	-37.11	-37.67	0.56	46.55	28.45	.99	7950	.243	826
		SD	17.22	16.03	4.93	2.94	3.86	.03	4395	.140	146
	Male	Mean	-25.13	-28.07	2.93	45.56	27.75	.96	8617	.198	738
		SD	18.59	14.52	7.92	2.53	2.62	.03	4483	.064	240
	Total	Mean	-29.63	-31.67	2.04	45.96	28.04	.97	8345	.219	779
		SD	18.67	15.50	6.93	2.70	3.13	.03	4374	.105	203
3	Female	Mean	-28.27	-30.27	2.00	46.74	27.13	.97	7766	.210	796
		SD	18.84	18.09	7.41	3.76	3.24	.04	4135	.084	146
	Male	Mean	-28.50	-33.10	4.60	46.82	27.27	.94	8633	.195	760
		SD	12.50	12.76	7.11	3.12	3.13	.04	2179	.099	191
	Total	Mean	-28.34	-31.16	2.81	46.76	27.18	.96	8055	.206	785
		SD	16.90	16.45	7.30	3.52	3.16	.04	3589	.084	158
4	Female	Mean	-22.38	-25.62	3.25	45.94	27.19	.96	7356	.284	820
		SD	16.00	16.41	5.93	3.30	3.90	.02	2698	.098	163
	Male	Mean	-26.30	-29.40	3.10	47.50	28.00	.94	6516	.235	646
		SD	19.99	18.45	6.62	1.51	3.92	.03	2795	.105	100
	Total	Mean	-23.88	-27.08	3.19	46.54	27.50	.95	7033	.255	745
		SD	17.36	16.96	6.07	2.82	3.85	.02	2712	.100	162
5	Female	Mean	-17.90	-22.05	4.15	47.55	26.85	.96	6997	.182	881
		SD	21.72	19.77	8.36	3.90	2.66	.03	3472	.052	196
	Male	Mean	-12.13	-13.57	1.43	45.12	25.71	.96	6205	.241	731
		SD	14.28	16.13	6.16	3.15	3.75	.03	2238	.120	174
	Total	Mean	-14.81	-17.51	2.70	46.23	26.23	.96	6573	.210	799
		SD	18.13	18.20	7.30	3.68	3.31	.03	2870	.093	197
Total	Female	Mean	-24.82	-27.05	2.24	46.87	27.27	.97	7310	.220	832
		SD	20.00	18.78	7.00	3.43	3.21	.03	3508	.092	167
	Male	Mean	-21.03	-23.63	2.60	46.09	27.08	.95	7212	.227	727
		SD	17.46	17.04	6.89	3.02	3.47	.03	3178	.098	185
	Total	Mean	-23.12	-25.51	2.40	46.52	27.18	.96	7266	.223	784
		SD	18.93	18.04	6.93	3.26	3.32	.03	3351	.094	182

Field of Study 1-5 codes: 1= Fine Arts, 2= Humanities, 3= Social Sciences, 4= Biological Sciences, Medicine, 5= Physical Sciences, Engineering, Maths, Systems Science; AQ (Full)= Autism Spectrum Quotient overall score; AQSI= Social Interaction AQ subscore; AQDet= Details/Patterns AQ subscore; Benton= Benton Face Recognition Test, # correct out of 54; RMET= Reading the Mind in the Eyes Test, # correct out of 36; 2D:4D= 2nd-to-4th digit length ratio; EFT= Embedded Figures Test latency, in ms; Motion Threshold= dot motion coherence threshold; Go/no-go RT= reaction time on motor inhibition task, in ms

Table 2. Full sample Principal Components Analysis of AQ Subdomains

	Component	
	Social Interaction	Details/Patterns
Attention Switching	0.710	-0.080
Communication	0.831	0.002
Details/Patterns	-0.114	0.984
Imagination	0.538	0.119
Social Skills	0.835	0.124

2.4.2 Intercorrelations Between Empathizing- and Systemizing Traits, and Field of Study

Correlation coefficients for the RMET, EFT, motion coherence threshold, reaction time, field of study, 2D:4D, and AQ full-scale and composite scores (AQSI and AQDet), by sex, are summarized in Tables 3 and 4, respectively. Tail probabilities are two-sided, and were calculated after each variable was z-transformed. Correlations were not subjected to multiple comparison corrections because (a) each hypothesis was generated *a priori,* with each having an independent theoretical and empirical justification; (b) testing E/S theory by nature means testing for interrelatedness among many measures broadly grouped under empathizing and systemizing labels, so whilst *a priori* hypotheses involved individual correlations, evidence for and against E/S relationships was interpreted in terms of patterns, not through conclusions drawn solely from potentially anomalous individual effects. Importantly, previous studies of the E/S spectrum in the literature have taken this same exact approach to multiple comparisons of correlated data (e.g., Carroll & Chiew, 2006; Voracek & Dressler, 2006; Walter, Dassonville, & Bochler, 2008) and, as such, direct comparisons to findings in the literature require such an approach.

Table 3. Female intercorrelation coefficients (Pearson's r)

	AQ (Full)	AQSI	AQDet	Benton	RMET	2D:4D	EFT RT	Motion Threshold	Go/no-go RT
Field	0.206	0.125	**0.255.**	0.000	-0.092	**-0.230.**	0.011	-0.193	0.123
AQ (Full)		**0.937..**	**0.344..**	-0.041	-0.045	-0.113	-0.189	0.159	0.061
AQSI			-0.006	-0.031	-0.060	-0.049	-0.202	0.163	0.035
AQDet				-0.034	0.032	-0.195	-0.004	0.004	0.082
Benton					0.152	-0.034	0.023	-0.213	0.080
RMET						-0.013	0.002	0.197	-0.113
2D:4D							0.057	-0.078	-0.034
EFT RT								0.000	-0.038
Motion Threshold									-0.243

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4. Male intercorrelation coefficients (Pearson's r)

	AQ (Full)	AQSI	AQDet	Benton	RMET	2D:4D	EFT RT	Motion Threshold	Go/no-go RT
Field	**0.307.**	**0.347..**	-0.081	-0.163	**-0.291.**	0.058	-0.184	0.083	-0.097
AQ (Full)		**0.920..**	**0.258.**	**-0.315.**	**-0.319.**	0.132	-0.035	-0.058	0.167
AQSI			-0.140	**-0.347..**	**-0.385..**	0.146	-0.059	0.056	0.216
AQDet				0.060	0.142	-0.026	0.056	-0.310	-0.123
Benton					**0.321..**	0.078	-0.146	0.131	-0.212
RMET						-0.153	-0.181	-0.035	0.072
2D:4D							0.095	-0.109	-0.093
EFT RT								0.123	0.084
Motion Threshold									-0.014

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

In women, higher-systemizing fields were significantly related to the AQ Details composite score (AQDet), $r(76) = 0.255$, p = .026, as well as 2D:4D, $r(78) = -0.230$, p = .042; in men, these correlations were much lower or non-existent. Conversely, men's fields of study were significantly related to higher Full-scale AQ scores, $r(62) = .307$, p = .015, AQSI composite scores, $r(62) = 0.347$, p = .006, lower RMET scores, $r(65) = -0.291$, p = .019, as well as lower Benton scores, though non-significantly, $r(65) = -0.163$, p = .19; each of these correlations were much lower or non-existent in women. Higher AQSI composite scores were also related to significantly lower Benton scores, $r(62) = -0.347$, p = .006, and RMET scores, $r(62) = -0.385$, p = .002, in men; in contrast, neither of these relationships were present in women. Motion

coherence threshold was related to lower AQDet subscores in men, though marginally non-significantly (perhaps due to the small number of available motion coherence thresholds), $r(27) = -0.310$, $p = .12$, whereas in women this correlation was essentially zero. Prolonged go/no-go response time was related to higher AQSI subscores in men, though marginally non-significantly, $r(55) = -0.216$, $p = .11$, whereas in women there was no association between reaction time and either AQ composite score. Amongst the cognitive measures, men's, but not women's, RMET and Benton scores were significantly, positively related, $r(65) = 0.321$, $p = .009$.

2.4.3 Post-hoc AQ Subdomain Principal Components Analyses, Separately By Sex

Based on above correlations trending toward sex-dependent covariance structures in empathizing and systemizing traits, we repeated the AQ subdomain score PCA to see if there were a similar sex difference in factor loadings, with more overlap between social and non-social traits in men than in women. Factor loadings were essentially unchanged from the original PCA for women (Table 5); Attention Switching, Communication, Imagination, and Social Skills subdomains loaded on a Social Interaction component (accounting for 46% of the total variance), whilst the Details subdomain comprised the majority of the second component, which accounted for 20.7% of the total variance. By contrast, much more overlap between subdomains was evident in men (Table 6); while the first component was similar to the Social Interaction factor from the full sample PCA (accounting for 40.1% of the total variance), the second component (accounting for 21.3% of variance) included the Details subdomain, but also had positive Communication and Social Skills loadings.

Table 5. Female Principal Components Analysis of AQ Subdomains

	Component	
	1	2
Attention Switching	.760	.089
Communication	.846	-.209
Details/Patterns	-.020	.984
Imagination	.543	.122
Social Skills	.854	.073

Table 6. Male Principal Components Analysis of AQ Subdomains

	Component	
	1	2
Attention Switching	0.648	-0.372
Communication	0.789	0.321
Details/Patterns	-0.260	0.880
Imagination	0.524	0.082
Social Skills	0.816	0.213

2.4.4 Group Differences

A series of 2 (Sex) x 5 (Field of Study) analyses of variance was performed to test for main effects of Sex and Field, and Sex-by-Field interactions, for each dependent measure. A main effect of Field arose for AQ scores, $F(4, 128)= 3.773$, $p = .006$, and *post hoc* t-tests using Tukey's method revealed a difference between high systemizing fields and both social sciences and humanities, both $p = .016$. Main effects for Sex arose for 2D:4D, $F(1, 133)= 8.409$, $p = .004$ (with men's digit ratios being lower on average than women's, as expected), and motor inhibition, $F(1, 119)= 7.76$, $p = .006$, as men's go/no-go reaction times were greater than women's, on average.

2.5 Discussion

In this sample of normal young adults, the relationship between empathizing and systemizing was highly sex-dependent. In men, empathizing and systemizing, or social and non-social "autistic" traits, were inversely related, indicating a mainly unidimensional scale of male empathizing and systemizing abilities. In women, empathizing and systemizing were largely independent, implying a multi-dimensional factor structure of autistic traits. Notable exceptions to these generalizations existed for each sex, though.

In men, being in a systemizing field of study was associated with "autistic" social difficulties measured by the four AQ subscales comprising the AQ Social Interaction (AQSI)

component (AQSS, AQComm, AQImag, AQAttSw) that was also related to decreased face recognition ability, and to decreased RMET scores; similarly, field was inversely related to these two face processing measures, with higher systemizing fields associating with lower Benton and RMET scores. At the same time, field was unrelated to traits typically associated with systemizing ability, such as motion coherence threshold, EFT latency, and the AQDet subscale. Thus, for men being in a systemizing field was related to poor empathizing and, at least within this study's set of measures, was more closely correlated with poor empathizing than with high systemizing. Similarly supporting an inverse E/S relationship in men, increased reaction time on the go/no-go measure of motor inhibition was directly related to the AQSI (though not significantly so), suggesting that inhibition of complex social behaviors may bear some relation to inhibition of simple motor behaviors; men in the current study, it follows, incurred reaction time costs of inhibiting the impulse to repetitively "shoot" at every ship emerging from a wormhole.

In women, fewer significant associations between measures arose overall, both within and between empathizing and systemizing domains. The few associations that did arise were within the latter domain; being in a systemizing field was associated with elevated AQDet scores (a composite measure of the systemizing tendencies of detail-oriented cognition and pattern recognition) in women; systemizing fields in women were also related to masculine (lower) digit ratios. Also in contrast to men, women's choice of field was not related to the social interaction AQ subscale, or to either face processing measure. In men, field was more closely related to (poor) empathizing than systemizing strengths. In addition, if digit ratio is indeed a reliable index of prenatal testosterone exposure, then the predictions of EMB (ironically) applied only to females – perhaps because of a ceiling effect in males' testosterone exposure. Further supporting

female E/S independence, in women neither of the face processing measures was related to either of the WCC measures, motion coherence threshold and EFT latency.

Women exhibited a covariance pattern opposite men's in another sense; not only were women's empathizing and systemizing abilities largely independent, but so were abilities typically grouped under the empathizing label. This was not the case in men, where empathizing abilities were related to systemizing and to other empathizing measures (Benton and RMET, e.g.), and it may mean that, in females, the E/S covariance structure may be multi-dimensional in a way that cannot be accounted for by any current E/S model. Social cognitive tasks such as face recognition, emotion/intention reading, and communication, for instance, may each employ a more domain-specific strategy in women than in men.

On the male side of this argument are findings that indicate more extensive unidimensionality in men than can be accounted for in the single-axis E/S spectrum model. Contrary to expectations, but supporting this male E/S unidimensionality, men's RMET and Benton scores – which were positively related – were both *inversely* related to EFT latency (though non-significantly), meaning that faster disembedding was associated with *increased* face processing abilities; similarly, men's RMET scores were directly (but non-significantly) related to AQDet scores. Also contrary to expectations, in an exploratory analysis men's motion coherence threshold seemed *inversely* related to AQDet subscores. Rather than indicating a *positive* relationship between empathizing and systemizing in men, as was argued by Voracek and Dressler (2006) when they found positive correlations between males' EQ and SQ, and SQ and RMET scores, we believe our results make a case for the idea that males may apply systemizing skills, such as disembedding, to empathizing tasks such as face processing – by disembedding individual facial features relevant to face recognition and emotion/intention

reading.

These data may help elucidate two important questions echoed by Jarrold et al. about the causal and functional natures of any empathizing-systemizing links, questions that have concerned ASC theorists for some time (Frith & Happé, 1994; Baron-Cohen, 1995, e.g.). Jarrold et al. entertain the possibility that the apparent ToM/WCC link may be illusory, arising from the use of RMET/EFT correlations to test this hypothesis. While the RMET is purported to measure intention-reading, they say, part of the skill required to "read" mental states from faces is gestalt, or centrally coherent, visual perception. Taking this argument a step further to infer causality, and citing "other data" as evidence, they state that WCC likely leads to ToM deficits, not v*ice versa*, because "central coherence appears to be a lower level process than theory of mind", and "it appears highly unlikely that something as 'psychological' as theory of mind…could explain low-level visual bias effects" (p. 135). As for the *functional* link between ToM and WCC, the authors defer to Frith's (1989) conjecture that WCC makes integrating social information from numerous sources difficult, and raise the possibility that joint attention (the triangulation of attention between the self, an object, and the other, a ToM precursor) is the specific ToM skill affected by WCC.

Our results support a methodologically-based link between systemizing and face processing, as Jarrold et al. posit, but just as the correlation between RMET and EFT scores in our males was in the opposite direction to that of Jarrold et al., our methodological reasoning is likewise opposite theirs. In contrast to their argument that the EFT and RMET are linked due to the central coherence needed to accurately read intention from faces, we argue that men who are quicker at disembedding geometric figures may also be better at disembedding relevant facial features on the RMET and, perhaps, the Benton. In other words, these tests of empathizing,

when applied to the male population, may actually be measuring systemizing.

Despite our opposing findings and conclusions regarding the methodological link between the RMET and EFT, our results are in general consistent with the causal link between WCC and ToM proposed by Jarrold et al., though our model is slightly different. WCC has been conceptualized as an epiphenomenon of a more basic adaptive neural mechanism, a compensatory response to reduced attentional selectivity in ASC (Belmonte & Yurgelun-Todd, 2003). This same attentional issue hinders early joint attention and, some say, ToM (Charman, 2003). (Such a link between attention and social cognition is partly supported by our principal components analysis, wherein the attention switching subdomain of the AQ loaded on the first principal component with communication, imagination, and social skills, whilst attention to detail was the sole subdomain loading on the second component; this view is reinforced by the negative albeit non-significant correlation between AQAtt subscores and the RMET). The same compensatory cognitive style that leads to quicker figure disembedding, then, may also compensate for face processing difficulties by allowing men to systemize emotional and facial recognition through piecemeal processing of facial features. This ability to systemize empathy – and to get away with it without incurring significant social cognitive deficits – may distinguish the population of "extreme", yet normal, males from the autism-spectrum population. Normal females in our sample were actually closer to ASC in this sense, as disembedding was associated with more "autistic" communication and social skills traits, suggesting less ability to systemize social situations.

Sex-dependent E/S covariance structures notwithstanding, these data are generally inconsistent with the predictions of EMB and E/S theories. No sex differences arose on any measures except for 2D:4D ratio (a well established difference that is independent of EMB and

E/S theories) and go/no-go reaction time. This overall lack of sex differences is consistent with the lack of sex differences on the RMET and EFT found by Jarrold et al. (2000), but not with Voracek and Dressler's (2006, 423 subjects) or Carroll and Chiew's (2006, 48 subjects) findings of RMET and AQ sex differences (the latter study found no difference on the EFT, however). The absolute difference between male and female AQ scores in the current study exceeded that of Voracek and Dressler but was based on a sample less than a third as large, suggesting that an AQ sex difference may indeed exist but as a vanishingly small effect. The lack of RMET sex difference is more puzzling. One possibility is that our selective sampling method yielded a greater proportion of women in the highest systemizing fields and men in the highest empathizing fields than a truly random sample would have done; only men in the highest systemizing fields had lower RMET scores than women in the highest systemizing fields, whereas men in high empathizing fields had the highest scores of any group by more than a point, on average. Thus, male and female mean scores may have been slightly skewed toward one another due to artefacts of our particular sample.

These findings raise issues with the construct validity of AQ subdomains as distinct measures of empathizing and systemizing. On face, our replication of Hoekstra et al.'s PCA of AQ subdomains supports E/S theory. A fairly clean split was evident between the first and second principal components, which roughly translate into empathizing and systemizing, respectively. Additionally, the empathizing component includes attention switching (which, we have argued, may underlie empathizing deficits). However, our cognitive measures (EFT in females, Benton and RMET in males) correlate with only the first component. The second component is neither inversely related to Benton and RMET scores (as a single-axis E/S factor space would predict), nor inversely related to EFT response latencies (as both single- and dual-

axis models would predict) in either sex. This is perhaps because subdomains comprising the first component contain questions posed in social contexts, whereas none of the questions comprising the details/patterns subdomain are.

This social contextual distinction may help explain the aforementioned "clean split" between the two principal components along empathizing and systemizing lines. Explaining how the Social Interaction component predicts WCC does not require appeals to theoretical relationships between social and non-social domains, as the Social Interaction subdomains directly indicate WCC as much as they do social, attentional, and imaginative deficits. Two examples from the AQAtt and AQImag subdomains illustrate these explanations: "I can easily keep track of several different people's conversations", and "I find it easy to make up stories". While the first question does relate to attention switching ability, it is posed in a social context; likewise, making up stories requires some concept of how fictional characters would interact socially. Both these questions exemplify a social skill/WCC confound; keeping track of simultaneous conversations would be near impossible if one attempted to follow each word-for-word; instead, the "gist" or central coherence of each is extracted from the bits and pieces that reach awareness. As for the second question, stories, like conversations, are narratives with central coherence, not randomly assembled statements.

On the other hand, our post hoc AQ subdomain factor analyses revealed a sex difference in factor structure that is supported by our findings of empathizing and systemizing dependence in men, and independence in women. In men, there was overlap between subdomains; while the first component was similar to that of the full sample PCA, the second component included the Details subdomain, but also had positive Communication and Social Skills loadings. If the overlap between AQ subdomains were methodological rather than empirical, then one would

expect sex to be irrelevant to the amount of overlap between principal components; instead, more overlap was present for men than women, just as empathizing and systemizing were generally more related in men than in women according to our cognitive measures as well.

2.6 Conclusion

Empathizing and systemizing are largely independent in women, but largely dependent in men. Systemizing ability was coupled with sacrifices in social abilities in men but not in women in our sample, except insofar as men supplemented empathizing with systemizing, utilizing their disembedding skills to process faces. The flip side of this conclusion is that males who are poor systemizers may also be poor at some empathizing abilities.

Moreover, when females enter science, they seem to do so as much when they are good systemizers as when they are poor empathizers, whereas – at least according to this study's set of measures – men seem to enter science more specifically when they have empathizing difficulties. Despite the relationships between men's weaker empathizing skills and their stronger systemizing skills, when it comes to selecting an occupation men's choices may be determined more by their weaknesses than by their strengths: men's choice of a scientific field of study is associated with weakness in empathizing skills as indexed by the AQ Communication, Social, and Imagination subscales. All these relationships of course, exist at the level of populations, and all are correlational, and so it would be a fallacy and an injustice to use them to make proclamations about individuals and/or causal directions.

2.7 Relevance to Work Two

Within E-S theory's assumption of social/cognitive co-variance lies the secondary assumption that such E-S co-variance applies indiscriminately to both 'male' and 'female' ends of the E-S spectrum, with females exhibiting the high empathizing-low systemizing counterpart

to the low empathizing-high systemizing pattern characterizing males and ASC individuals. The broader importance of this study with regard to the present thesis, then, lies in the *sex-dependent* nature of the E-S co-variance patterns that arose in this experiment. Thus, contrary to the framing of the E-S co-variance hypothesis as a unilateral question of evidence for/against, E-S co-variance may instead be greater or more likely at the male/autistic than at the female end of the E-S spectrum.

In the next chapter, the ideas about sex differences in empathizing, systemizing, E-S co-variance, and field of study explored in this chapter are examined from the perspective of underlying biology and its downstream developmental effects. Specifically, we review evidence related to prenatal hormonal influences on neuroarchitectural organization, and whether such influences can account for the downstream sex difference patterns seen typically developing adults in the previous chapter; more specifically, we focus on evidence from studies utilizing an increasingly popular proxy measure of prenatal androgen exposure, 2^{nd}-to-4^{th} digit length ratio. Based on this body of evidence we argue, among other things, that the sex-dependent E-S Co-variance seen in the previous work may be rooted in sexually dimorphic timescales of androgen-related neuroarchitectural organization, and sex-dependent developmental trajectories of E-S Co-variance that result from such sexually dimorphic timescales.

REFERENCES

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the autism spectrum quotient (AQ). *Personality and Individual Differences, 38,* 451–460.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, Massachusetts: MIT Press/Bradford Books.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences,* 6, 248–254.

Baron-Cohen, S., & Hammer, J. (1997a). Is autism an extreme form of the "male brain"? *Advances in Infancy Research, 11,* 193–217.

Baron-Cohen, S., & Hammer, J. (1997b). Parents of children with Asperger syndrome: What is the cognitive phenotype? *Journal of Cognitive Neuroscience, 9,* 548-554.

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

Baron-Cohen, S., Wheelwright, S., Hill, J, Raste, Y, & Plumb, I. (2001a). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241-251.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. G., & Clubley, E. (2001b). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males, females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.

Belmonte, M.K., & Yurgelun-Todd, D. (2003). Functional anatomy of impaired selective attention and compensatory processing in autism. *Cognitive Brain Research, 17,* 651-664.

Benton, A.L., Hamsher, K, D., Varney, N.R. and Spreen, O. (1983). *Contributions to Neuropsychological Assessment: A Clinical Manual.* New York: Oxford University Press.

Carroll, J.M., & Chiew, K.Y. (2006). Sex and discipline differences in empathizing, systemizing and autistic symptomology: evidence from a student population. *Journal of Autism and Developmental Disorders, 36,* 949-957.

Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., Banerjee-Basu, S., & Baron-Cohen, S. (2009). Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger Syndrome. *Autism Research*, *2*, 157-177.

Dawson, G., Webb, S.J., Schellenberg, G.D., Dager, S, Friedman, S., Aylward, E., & Richards, T. (2002). Defining the broader phenotype of autism: genetic, brain, and behavioral perspectives. *Development and Psychopathology, 14,* 581-611.

Falter, C. M., Arroyo, M., & Davis, G. J. (2006). Testosterone: activation or organization of spatial cognition? *Biological Psychology, 73,* 132–140.

Frith, U. (1989). *Autism: Explaining the enigma.* Oxford, UK: Blackwell.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition, 50,* 115-132.

Geschwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization: Biological mechanisms, associations, and pathology.* Cambridge, Massachusetts: MIT Press.

Happé, F., Frith, U., & Briskman, J. (2001). Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: I. experimental tests. *J. Child Psychology and Psychiatry, 42,* 299-307.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9,* 1218-1220.

Hoekstra, R.A., Bartels, M., Cath, D.C., & Boomsma, D.I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders, 38,* 1555-1566.

Jarrold, C., Butler, D. W., Cottington, E. M., & Jimenez, F. (2000). Linking theory of mind and central coherence bias in autism and in the general population. *Developmental Psychology 36,* 126–138.

Kempel, P., Gohlke, B., Klempau, J., Zinsberger, P., Reuter, M., Hennig, J., (2005). Second-to-fourth digit length, testosterone and spatial ability. *Intelligence, 33,* 215–230.

Lord, C., Cook, E.H., Leventhal, B.L., & Amaral, D.G. (2000). Autism spectrum disorders. *Neuron, 28,* 355-363.

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

Manning, J. T., Baron-Cohen, S., Wheelwright, S., & Sanders, G. (2001). The 2^{nd} to 4^{th} digit ratio and autism. *Developmental Medicine and Child Neurology, 43,* 160-164.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 32,* 1081-1105.

Pentland, A. P. Maximum likelihood estimation: the best PEST. (1980). *Perception and Psychophysics, 28*, 377-379.

Piven, J. (1999). Genetic liability for autism: the behavioral expression in relatives. *International Review of Psychiatry, 11,* 299-308.

Romano, M., Leoni, B., & Saino, N. (2006). Examination marks of male university students positively correlate with finger length ratios (2D:4D). *Biological Psychology, 7,* 175-182.

Voracek, M. & Dressler, S. (2006). Lack of correlation between digit ratio (2D4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemizing, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41,* 1481-1491.

Walter, E., Dassonville, P., & Bochsler, T. M. (2008). A specific autistic trait that modulates visuospatial illusion susceptibility. *Journal of Autism and Developmental Disorders, 39,* 339-349.

Witelson, S., & Nowakowski, R. (1991). Left out axons make men right: A hypothesis for the origin of handedness and functional asymmetry. *Neuropsychologia, 29,* 327–333.

Witkin, H. A., Oltman, P., Raskin, E., & Karp, S. (1971). *A manual for the Embedded Figures Tests.* Palo Alto, CA: Consulting Psychologists Press, Inc.

CHAPTER 3

BRAIN ORGANIZATION THEORY, DIGIT RATIO (2D/4D), AND SEX DIFFERENCES IN PREFERENCES AND COGNITION

3.1 Introduction

Women's growth in the scientific workforce has been meteoric over the past 40 years, with increases of several orders of magnitude (Ceci & Williams, 2010a,b). For example, in 1970 fewer than 5% of scientific and medical doctorates were awarded to women, but by 2006 approximately 50% of MDs and 48-51% of PhDs in biology were being awarded to women, as were 76% of doctorates in veterinary medicine, and 67% of PhDs in psychology.

However, there is one glaring exception to women's progress in scientific careers. In fields that are highly quantitative, women's success has been far less pronounced. In 2006 less than a third of the PhDs in highly quantitative fields were awarded to women: mathematics-29.6%, computer science-21.3%, physical sciences-29%, engineering-20.2% (Burrelli, 2008, Table 1). In the top 100 U.S. universities, only 8.8% (mechanical engineering) to 15.8 (astronomy) % of all professorial ranks combined in many quantitative fields are occupied by women (Nelson & Brammer, 2010, Table 11). Among full professors, women usually number < 10%: chemistry—9.7%, mathematics—7.1%, computer science—10.3%, physics—6.1%, chemical engineering—7.3%, civil engineering—7.1%, electrical engineering—5.7%, mechanical engineering—4.4%, and economics—8.7%. The reason for women's lower representation in these fields has been the source of heated debate (Ceci & Williams, in press; Shalala, Agogino, Bailyn, Birgeneau, Cauce, Deangeles, et al., 2007; Sommers, 2008) and, like many other controversies in the social policy realm (e.g., violence, addiction, sexual orientation), has been framed in terms of the relative influence of nature versus nurture, although recently it

has been recast in terms of a biopsychological vs. main effects view of sex differences in which biological factors are enmeshed with social forces in an iterative unfolding (Berenbaum & Resnick, 2007; Bronfenbrenner & Ceci, 1994; Guo & Stearns, 2002; Halpern, Benbow, Geary, Gur, Hyde, & Gernsbacher, 2007).

To be sure, neither side contends a mutual exclusivity of nature or nurture in explaining the relatively slow progress made by women in math-intensive fields, but their emphases are clearly at opposite poles (Ceci & Williams, 2007). One side emphasizes sociocultural factors as primary, such as early socialization practices, and biased teachers' and parents' attitudes toward girls and mathematics, with a resultant stereotype threat and discriminatory hiring and promotion practices, and "chilly work climate". This side deemphasizes biological sex differences in ability (Shalala et al., 2007; Spelke, 2005), and points to rapid changes in the proportions of females scoring at the extreme right tail of the mathematics distribution (i.e., those scoring in the top .01% or 1 in 10,000), from 13:1 in the early 1980s to 4:1 by the mid-1990s, where it has remained since (Wai, Cacchio, Putallaz, & Makel, 2010). Indeed, such dramatic changes over such a brief time period are inconsistent with mathematical ability tightly controlled by biology, but more consistent with Hyde's *Gender Similarity Hypothesis* (Hyde, 2005; Else-Quest, Hyde, & Linn, 2010). Additionally, proponents of sociocultural explanations point to the large variance in cross-cultural analyses, with females outperforming males in some countries (see Ceci, Williams, & Barnett, 2009; Else-Quest et al., 2010), and the best predictor of international sex differences is the degree to which its citizens exhibit implicit gender-science stereotypes (Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., et al. 2009).

In contrast, those who assign a greater role to biology stress hormonal, neural, and genetic factors that are alleged to result in a male advantage in spatial and mathematical abilities (for

review of genetics and hormone findings, see Ceci et al., 2009; Ceci & Williams, 2010). Also, logically, biology cannot be ruled out on the basis of temporal changes or transnational differences because even highly heritable characteristics such as height are sensitive to environmental changes but the existence of these changes is still compatible with high heritabilities (Ceci, 1996). Due to greater male variability that is alleged to be biologically-based, the farther out on either tail, the higher the ratio of males to females. (Male variance in mathematical and spatial performance is with some exceptions 10-20% greater than female variance, reflected in greater asymmetries at the tails of the distribution. Notwithstanding greater male variability in most studies, this sex difference in variability is not ubiquitous and can vary with culture and gender equality attitudes—see Feingold, 1994; Hyde & Mertz, 2009; Penner, 2008.)

The discussion that follows focuses on the most popular biological account of women's underrepresentation in math-intensive fields of science—namely, the putative role of prenatal exposure to hormones such as sex steroids (the most important being testosterone) on early brain organization, which in turn leads to sex differences in spatial and mathematical abilities as well as to a plethora of other outcomes (personality, sexual orientation, interests). This sexually dimorphic brain organization hypothesis (e.g., Baron-Cohen, 2003), termed "brain organization theory", posits that male brain physiology is inherently "built" for efficient spatial and quantitative processing by these sex steroids. Sex steroids increase lateralization and a bias toward right hemispheric processing, the right hemisphere being the preferentially involved in spatial and numerical processes, particularly those dealing with abstract numerical relations (Baron-Cohen, 2003). It is also argued by advocates of the biological position that steroids render the male brain more sensitive to the activational effects of testosterone. This may result

not only in better ability to undertake 3-dimensional spatial rotation, but also in a host of behavioral changes such as higher risk taking, search persistence, and heightened vigilance and faster reaction times. Meanwhile, inherently female brain physiology is said to be "built" for language fluency, empathy, emotion recognition, and other processes that involve more inter-hemispheric coordination, having reduced lateralization and an overall bias toward left hemispheric processing.

We begin with a précis of the evidence for sex differences in mathematical and spatial ability. Following this, we describe the dominant biological and sociocultural arguments put forward to account for them, focusing especially on claims that *in utero* hormone exposure organizes brains differently in males and females, leading to male advantages in math and spatial cognition. These claims of prenatal hormone influence are then evaluated in the context of studies (summarized in Table 7) relying on a putative biomarker of prenatal hormone exposure, the length ratio of the index (second) to ring (fourth) finger, or 2D/4D ratio. We argue that despite a wealth of 2D/4D-based evidence supporting brain organization theory, there remains very significant inconsistencies and contradictions in the 2D/4D literature. We conclude by arguing that these issues mirror the larger issues plaguing brain organization theory as a whole; that despite evidence of the role of early brain organization in women's underrepresentation in mathematically-intensive careers, there remains very significant inconsistencies and contradictions. We then point out numerous alternative explanations of key findings that will have to be resolved for early brain organization theory to be a compelling account of sex differences across cognitive, social, behavioral, and sexual domains.

3.2 Evidence of Sex Differences in Mathematics and Spatial Ability

Hundreds of published studies have reported sex differences in mathematical and spatial

ability. The general pattern of findings is one of overall male superiority when it comes to extreme ability at the right tail of the distribution; but the size of the right tail sex differences vary by sample and particularly by culture (Feingold 1994; Hyde & Mertz, 2009). In addition, extreme right tail sex differences can be contrasted with mean performance, where sex differences are small or non-existent. Effect sizes for male superiority in 3-dimensional mental rotation—an ability that has been circumstantially linked to some forms of mathematical prowess—usually fall in the moderate to large range (e.g., in a number of large meta-analyses, *d*s ~ 0.50 to 0.80). Illustrative are the various meta-analyses reported by Hyde and her colleagues. Hyde (2005) synthesized 128 effect sizes on a broad range of measures from 47 published meta-analyses and reported large effects for mental rotation and mechanical reasoning favoring males (*d*s between .56 and .76). Other forms of spatial ability, however, do not exhibit consistent male superiority, such as rotation of 2-dimensional figures, and still others tend to be associated with female superiority, such as spatial memory. So while there is a general male superiority in the extreme right tail from which the professorial scientific workforce is drawn, caveats such as distinguishing between extreme right tail and mean sex difference, and inconsistencies such as the 3-D male advantage but not a 2-D advantage, are recurring themes in the sex difference literature.

Along the same lines, hundreds of studies document male superiority for mathematical ability—if that is defined as a score on one of the national or international aptitude tests such as the Third International Mathematics and Science Study (TIMSS), the Program for International Student Assessment (PISA), the National Assessment of Educational Progress (NAEP), or the SAT-Mathematics (SAT-M). Again, there are many caveats, most notably that early mathematical sex differences are nil or even favor females (Ceci et al., 2009). This has been

attributed by some to the role of the left hemisphere in learning basic arithmetic concepts (Brosnan, 2006). Young girls' sometimes superior performance in math, in other words, is purportedly influenced by their superior ability to learn and remember things verbally, and is thus more a matter of memorization and verbal ability than inborn mathematical ability (though see later argument against a verbal account of the female grade advantage in mathematics courses). However, beginning around middle school, the effect sizes for male superiority appear and increase throughout high school. To some, this trend suggests a developmental convergence of a sex difference in sensitivity to activational effects of testosterone, determined prenatally, and increased levels of testosterone in males during this middle school-age, pubertal period.

A related source of the inconsistency is the nature of the mathematics being tested at different ages, with the more abstract and complex math being associated the largest effect sizes in favor of males. In fact, at any given age, the items that are most difficult are associated with the largest male superiority (e.g., Ceci et al., 2009; OECD, 2004). Relatedly, sex differences in mathematics depend critically on not only the type of math being tested and at what age, but also on where in the score distribution one looks: As noted, there are no systematic sex differences at the midpoint of the distribution, but fairly large differences at the extreme tails, both right and left. Among the top 1% of students on standardized mathematics tests, there are approximately 2 males for every female, and this ratio has been found across a wide variety of nationally representative samples (e.g., Hyde, Lindberg, Linn, Ellis, & Williams, 2008, n=*7 million US students in various grades*; Lohman & Lakin, 2009, n=318,599 9-11-year-olds; Strand et al., 2006, n=320,000 11-year-olds; Mullis, Martin, Fierros, Goldberg, & Stemler, 2000, n= .5 million; Wai et al., 2010, n=1.6 million 7th graders). Because overall male score variability is roughly .15 SD greater than females', the farther out on either tail, the higher the ratio of males

to females. During the 2006-2010 period, in a non-random sample of 7th grade perfect-800 scorers on the SAT-M, there were 6.58 males for every female (Wai et al., 2010). In the past 20 years, there have been 37 perfect scorers among 7th graders on the ACT-Science test, 36 of whom were male.

A final source of inconsistency revolves around the achievement-aptitude distinction. Females get better grades in mathematics classes than males throughout high school and college (Gallagher & Kaufman, 2005); males outperform females at the right tail of aptitude tests such as the SAT-M, the GRE-Q, and the NAEP. This has led to claims and counterclaims of bias, with one side arguing that aptitude measures such as the SAT-M underpredict female performance in math classes, and the other side arguing that classroom teachers are biased against males because they receive grades below what are predicted from their aptitude scores (Ceci et al., 2009). For example, in a study of 47,000 college mathematics students, of whom 14,019 males and 10,087 females took calculus, males who received grades of D and F had SAT-M scores equal to women who received grades of B. The average SAT-M gap between males and females in calculus was 38 points (Wainer & Steinberg, 1992, Table 3, column 2). The difficulty in distinguishing between achievement and aptitude underpins much of the controversy. As noted by others (e.g., Ceci, 1996; Halpern et al., 2007), it is both conceptually and empirically difficult to measure ability without also measuring achievement to some extent, leading some to *"use the term ability as it was defined by Fleishman (1972): a general trait of an individual that is the product of learning and development."* (Halpern et al., 2007, p. 3).

3.3 The Putatively Causal Role of Brain Organization Theory in Math and Spatial Performance

To recap, ample evidence exists that, notwithstanding important caveats about where in

the score distribution, and at what age, and on what types of tests are considered, males generally outperform females on math and spatial tasks linked to math achievement, and the effect sizes of these differences are substantial. Moreover, some take these sex differences in math and spatial ability as part of a constellation of sex differences that support a particular nature-based theory, the so-called brain organization theory, in which hormones, particularly prenatal exposure to male hormones, are said to induce sexually dimorphic brain organization patterns that manifest in personality, toy preferences, sexual orientation, activity level, aggression, cognitive profile, and sex differences in education and occupations (for review of the myriad manifestations of male hormones, see Jordan-Young, 2010).

According to *callosal theory* (Witelson & Nowakowski, 1991), prenatal testosterone mediates early axon pruning in callosal tissue, and thus the more testosterone a brain is exposed to *in utero*, the more lateralization there is; evidence of less lateralization in females supports this assertion (Wiesniewski, 1998). In addition, prenatal testosterone is alleged to encourage right hemisphere growth, while simultaneously slowing left hemisphere growth (Geschwind & Galaburda, 1987). In sum, prenatal level of testosterone is viewed as the X factor in the relationship between sex and an androgen-induced brain organization well-suited for mathematical and spatial problem solving. For the remainder of the manuscript we focus on a particular, morphologically-based body of evidence used with increasing frequency, usually in support of the argument that sex differences in mathematical and spatial performance are influenced by prenatal brain androgenization.

3.4 Testing Brain Organization Theory: 2D/4D as a Proxy for Prenatal Testosterone Levels

Although the cognitive effects of prenatal hormones are unclear, these same hormones do without question affect the growth of body regions, including the symmetry of limbs and digits,

as well as the urogentialia, during gestation. One such effect is the length ratio of the index (second) finger vis-à-vis ring (fourth) finger. Since increasing levels of prenatal testosterone exposure lead to greater relative length in the fourth digit compared to the second digit (Manning, 2002), the 2D/4D ratio is viewed by many as a proxy for the relative level of prenatal testosterone exposure *in utero*. Because the brain organization patterns outlined by brain organization theory arise around the same time as digit length determination (during the first trimester or early in the second trimester of pregnancy), 2D/4D ratio, the logic goes, provides a rough idea of the amount of lateralization occurring during this organizational period (Brosnan, 2006). More specifically, homeobox genes (HOXA and HOXD in particular), activated around the 14th week of gestation and responsible for differentiation of both gonads and digit growth (Kondo, Zakany, Innis, & Duboule, 1997), are viewed as the basis of the prenatal testosterone-digit ratio link.

Convergent validation of 2D/4D as a proxy measure of relative hormone levels comes from findings indicating that digit ratio is correlated with the testosterone-to-estrogen ratio found in amniotic fluid (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004). Lower 2D/4D ratios are characteristically male (0.98, on average, compared to the female average of 1), a result of greater prenatal testosterone. Moreover, low 2D/4D in males is associated with many of the same traits typically associated with current (as opposed to prenatal) testosterone levels, such as reproductive success (Manning, Barley, Walton, Lewis-Jones, Trivers, Singh et al., 2000) and high sperm counts, but also cognitive ability and left hand preference. Similarly, in women 2D/4D is positively related to estrogen levels, and a low (male-like) 2D/4D is predictive of female homosexuality (Okten, et al., 2002). 2D/4D even predicts age of onset of breast cancer in women (Manning & Leinster, 2001).

The 2D/4D ratio is relatively stable from 2 years of age onward, and circulatory (current) testosterone levels are unrelated to this ratio, meaning that 2D/4D is an indication of prenatal hormone levels specifically, rather than hormone levels in general (Hönekopp, Bartholdt, Beier, & Liebert, 2007). Thus, 2D/4D ratio is now commonly used as a reliable proxy measure of prenatal hormone levels.

3.5 Evidence for Biologically-Based Cognitive Sex Differences From 2D/4D Studies

Taking evidence in the literature at face value, 2D/4D studies do generally support Brain Organization Theory. Although most of the significant 2D/4D-psychometric correlations in the literature have small effect sizes, as can be seen in Table 7 the magnitude of several effect sizes are in the moderate to large range (e.g., Brosnan 2006, 2008; Table 7). Upon closer inspection, however, the evidence for a link between prenatal hormones and spatial and numerical abilities provided by studies using the 2D/4D ratio is as equivocal as the cognitive sex difference findings reviewed above, in three ways outlined briefly here and elaborated in detail below. First, while some studies indicate the linear relationship between 2D/4D ratio and spatial/mathematical ability that is predicted by Brain Organization Theory, other studies indicate an inverse-U relationship between 2D/4D and spatial (specifically 3-D mental rotation) and numerical abilities. In the latter studies, optimal levels of prenatal testosterone for spatial abilities fall in the low-male/high-female range. Second, whether the relationship between the 2D/4D ratio and spatial/ mathematical abilities is quadratic or linear, some studies indicate that prenatal brain organization pathways and/or timing of those pathways may be sexually differentiated. This is because some 2D/4D studies have alternately indicated that 2D/4D ratio predicts spatial/numerical abilities in males but not in females, or in females but not in males.

Table 7. Summary of Cognition-Related 2D/4D Studies Reviewed

	Sample characteristics			2D/4D *M* (*SD*)				
Study	Age range and/or *M* (*SD*)	*N* (males, females)	Nationality	Males	Females	Main findings	Effect size(s) of main findings	Supports brain organization theory?
Alexander (2006)	18–22y	64 35m, 29f	American	0.947 (0.03)	0.965 (0.03)	Participants with visual fixation bias toward male-typical toys had lower 2D/4D than those with visual fixation bias toward female-typical toys, in males (1) and females (2).	(1) $d = 0.56$ (2) $d = 0.63$	Yes
Brookes, Neave, Hamilton, & Fink (2007)	Males: $M = 24.4$y Females: $M = 22.5$y	80 40m, 40f	British (mainly Caucasian)	RH: 0.965 (0.03) LH: 0.954 (0.03)	RH: 0.979 (0.03) LH: 0.970 (0.02)	Females with below-median 2D/4D had greater functional lateralization than females with above-median 2D/4D (1). Males with above-median 2D/4D had greater functional lateralization than males with below-median 2D/4D (*n.s.*).	(1) d = 0.46	Yes (females) and No (males)
Brosnan (2006)	23–62y $M = 44$y (10)	107 83m, 24f	British (mainly Caucasian)	0.987 (0.03)	0.984 (0.02)	Faculty from math-intensive disciplines had higher 2D/4D than social science faculty.	$d = -0.63$	No
Brosnan (2008)	6–7y	75 33m, 42f	British (mainly Caucasian)	0.950 (0.03)	0.960 (0.03)	Full sample 2D/4D negatively correlated with within-subject differences between math and verbal abilities (1). Males' 2D/4D negatively correlated with math ability (2). Females' 2D/4D positively correlated with verbal ability (3).	(1) $d = 0.45$ (2) $d = 0.75$ (3) $d = 0.54$	Yes
Bull & Benson (2006)	18–52y $M = 22.9$y (7.4)	75 37m, 38f	British	(Medians) RH: 0.964 LH: 0.963	(Medians) RH: 0.979 LH: 0.985	Significant correspondence between mental and spatial numerical representations present in RH (1) and LH (2) of below-median 2D/4D group, independent of sex. No effects for above-median 2D/4D group.	(1) $d = 0.23$ (2) $d = 0.19$	Yes
Coates, Gurnell, & Rustichini (2009)	19–38y $M = 26.9$y (4.1)	48 48m, 0f	British (mainly Caucasian)	0.959 (.004)	*N/A*	2D/4D negatively correlated with overall profitability in high frequency stock traders.	$d = 1.1$	Yes
Falter, Arroyo, & Davis (2006)	Males: 19–34y $M = 24.1$y (4.6) Females: 20-41y $M = 24.1$y (3)	46 24m, 22f	British (mainly Caucasian)	0.958 (.03)	0.970 (.03)	Curvilinear relationship between 2D/4D and targeting speed, independent of sex (1). Linear, negative relationship between 2D/4D and figure disembedding (EFT) speed, independent of sex (2). 2D/4D did not predict mental rotation.	(1) $d = -1.12$ (2) $d = 0.67$	Yes (2) and No (1)
Fink, Brookes, Neave, Manning, & Geary (2006)	6–11y $M = 9.3$y (1.3)	73 35m, 38f	Austrian, British (all Caucasian)	RH: 0.969 (0.03) LH: 0.962 (0.03)	RH: 0.994 (0.03) LH: 0.994 (0.03)	Males' 2D/4D negatively correlated with numerical ability (number knowledge, counting, visual representation) in RH (1) and LH (2). No effect in females (*n.s.*).	(1) $d = 1.18$ (2) $d = 0.76$	Yes (1, 2) and No (females)
Kempel, Gohlke, Klempau, Zinsberger, Reuter, & Hennig (2005)	Males: $M = 24.2$y (4.2) Females: $M = 23.5$y (4.3)	39 17m, 23f	German (mainly Caucasian)	0.963 (0.02)	0.983 (0.03)	Females with below-median 2D/4D outperformed those with above-median 2D/4D on numerical (1) and spatial (2) tests. No effects for males.	(1) $d = 0.48$ (2) $d = 0.68$	Yes (1, 2) and No (males)
Romano, Leoni, & Saino (2006)	21–25y $M = 22.8$y (.06)	204 84m, 124f	Italian (mainly Caucasian)	*N/A*	*N/A*	Male RH 2D/4D positively related to examination marks in biological and natural science college courses. No effects for male LH, female LH, and RH.	$d = 0.67$	No
van Anders & Hampson (2005)	18–42y $M = 23.8$y (5.7)	99 0m, 99f	Canadian (mainly Caucasian)	*N/A*	RH: 0.970 (*N/A*) LH: 0.975 (*N/A*)	2D/4D unrelated to any of three spatial tests (paper folding, mental rotation, spatial orientation) in all-female sample.	all *n.s.*	No
Valla et al. (2010)	18–28y $M = 20.2$y (2.3)	124 65m, 79f	American (mainly Caucasian)	0.950 (0.03)	0.970 (0.03)	2D/4D negatively correlated with math-intensiveness of college major in females, but not males.	$d = 0.47$	Yes (females) and No (males)
Weis, Firker, & Hennig (2007)	17–53y $M = 26.5$y (N/A)	47 26m, 21f	German (mainly Caucasian)	RH: 0.931 (0.05) LH: 0.952 (0.04)	RH: 0.962 (0.04) LH: 0.961 (0.03)	Males' RH 2D/4D negatively correlated with preference for "Realistic" careers (1). Females' LH 2D/4D negatively correlated with preference for "Investigative" careers (2).	(1) $d = 0.95$ (2) $d = 0.98$	Yes (1, 2)
Overall	$M = 23$y	1081 527m, 579f		0.961 (0.03)	0.974 (0.03)			

Note. Overall 2D/4D means might not reflect true population means, because demographic diversity was limited within and across samples. y = year; m = male; f= female.

Last, some 2D/4D studies indicate that prenatal hormone exposure may be equally (or more) responsible for inherent spatial and numerical *preferences* than for actual *ability* in these domains. In this scenario, androgenization of the brain may tilt early preferred play toward spatial activities (block-building, gross motor play) which may foster later spatial and numerical achievement, as some sex difference research indicates (Ceci et al., 2009). The implication of this last point is that the influence of prenatal testosterone on mathematical abilities is played out post-natally, and through experientially rather than innately-garnered abilities. This last possibility is particularly important in the broader context of underrepresentation of women in math-intensive fields, and may help reconcile the nature and nurture sides of this larger debate (e.g., Halpern, 2004; Shonkoff, Phillips, et al; National Research Council, 2000). This is because although it concedes a biologically based sex difference, it is a sex difference that is more amenable to counteraction via intervention than innate differences in ability might imply to some. In other words, if the real sex difference is one of *preference* for spatial play, and not *ability*, then ameliorating the problem of underrepresentation may be a matter of influencing girls' preferences before their biological potentials in math suffer from the snowballing effects of early preferences.

3.5.1 The Inverse-U Relationship Between the 2D/4D Ratio and Spatial/Numerical Abilities

The idea of an inverse-U relationship between prenatal testosterone exposure levels and spatial/mathematical ability is not new, nor is it confined to 2D/4D-based interpretations (see Brosnan, 2006). For instance, it has been hypothesized that, while increasing levels of testosterone may foster right hemisphere growth and accompanying spatial ability, too much testosterone could slow growth in both hemispheres, leading to decreases in spatial ability (Geschwind & Galaburda, 1987, cited in Brosnan, 2006).

Albeit speculative, some evidence from 2D/4D studies is consistent with this possibility. For instance, in one study 2D/4D ratio was significantly related to field of study among university faculty (Brosnan, 2006): science faculty members were found to have a more feminine (higher) 2D/4D ratio compared to social science/humanities/management faculty. Notwithstanding this intriguing finding, however, engineering faculty in this study had digit ratios between, and not significantly different than, those of science and social science/humanities/management faculty, despite engineering presumably requiring more spatial and math skills than many fields of science such as biology, save those subfields within it that are highly quantitative such as computational and population biology. A similar curvilinear relationship has been shown between salivary testosterone and spatial abilities such as 3-D mental rotation (Moffatt & Hampson, 1996), though as previously noted circulatory testosterone levels are distinct from prenatal (organizational) levels. A U-shaped, rather than inverse U-shaped, pattern has also characterized correlations between prenatal testosterone levels (measured via amniocentesis) and the amount of eye contact made with a parent (Lutchmaya, Baron-Cohen, & Raggatt, 2002). Such a relationship is said to typify the "female brain", which is touted as the converse of the purported relationship between testosterone and spatial/numerical ability. This "female brain" relationship is viewed as complementary evidence for a testosterone-spatial ability causal link. However, citing such a finding as support for an inverse-U relationship between hormones and spatial/numerical ability assumes not only direct relationships between preference and ability, and between *early* preference and *later* ability, but it also assumes that spatial/numerical and emotional/interpersonal abilities are inversely related, which remains an open question (Valla et al., 2010).

Yet there is still further evidence for an inverse-U relationship from other digit ratio

studies in the literature. For instance, studies have found a U-shaped relationship between targeting latency and 2D/4D ratio, such that extremely low and extremely high digit ratios were related to more time needed to accurately move a cursor to a target stimulus (Falter, Arroyo, & Davis, 2006). However, the same study found a direct, linear relationship between 2D/4D ratio and spatial ability as measured with a perceptual disembedding task. This calls into question the generality of the relationship and suggests that targeting entails specific operations that go beyond spatial, such as eye-hand coordination that may be influenced by playing dynamic videogames.

The numerical and mental rotation skills of females in another study appeared to benefit from a 2D/4D ratio in the low (more masculine) female range, while digit ratios in the high (more feminine) male range benefited males' numerical abilities (Kempel et al., 2005). Interestingly, however, no complementary relationships were found for verbal ability, which might be expected if prenatal testosterone diminishes left hemisphere growth (Geschwind & Galaburda, 1987). Lastly, male digit ratios have also correlated positively with performance in university science classes, with class performance related to more feminine digit ratios among male science students (Romano, Leoni, & Saino, 2006). However, the authors note that the positive correlation found for male students may have been an artifact of the oral and written nature of many of the course assignments and exams. Such assessments, the argument goes, favor verbal over spatial/numerical skills. However, this study was done with Italian university students in a system that uses verbal assessments of math and science far more heavily than is true of mathematical assessments in the U.S., which tend to require algebraic manipulation and the ability to draw graphs rather than language expression. This claim of female grade superiority as a result of verbal assessment procedures has been argued to explain away female

superiority in mathematical aptitude. (Ceci & Williams, 2010). However, such verbally-laden mathematics assessment is found more at younger ages where word problems predominate pre-algebra than among more advanced mathematics courses at which females continue to excel. Thus this body of evidence favoring the hormone-math/spatial/2D/4D linkage has been fraught with inconsistent findings and equivocal interpretations.

Evidence for a quadratic relationship between 2D/4D and spatial skills from investigations using *functional* lateralization has been similarly conflicting. In these studies, functional lateralization is inferred from response time differences between numerical recognition trials in which response keys were in the same versus opposite direction as the mental, analogical magnitude representation of numbers along a number line (left=lower magnitude, right=higher magnitude) This difference, also known as the SNARC (Spatial Numerical Association of Response Codes) may be preferable to using batteries of spatial/numerical tests for two reasons. First, the spatial analog of magnitude used in this paradigm is viewed by some as a metacognitive Rosetta Stone for transformations between various modes of numerical and spatial processing (auditory, visual Arabic, etc.). Second, and in terms of digit ratio studies in particular, functional lateralization may be a better intermediary between digit ratio and physical hemispheric lateralization than batteries of spatial/numerical tests. Thus SNARC evidence warrants special attention.

In one SNARC study, individuals with lower digit ratios exhibited greater degrees of functional lateralization than their high digit ratio counterparts, supporting the brain organization hypothesis (Bull & Benson, 2006). Importantly, however, this conclusion was not based on correlations, but rather on a comparison of “low” and “high” digit ratio distinctions based on median splits. The male median ratio was approximately 0.96, so even some individuals

categorized as having a "high" digit ratio may have had lower than average (which is 0.98 for males) ratios. At the same time, female ratios in this sample were low as well (median split of approx. 0.98), which may have balanced any male ratio biases in the analyses, which were performed on the full, mixed-sex sample. Thus the interpretation is ambiguous and could be the result of the sex distribution around the cut-off. (On a side note, sample representativeness is also an issue in Kempel et al. (2005) described above, particularly in their male sample, where the mean digit ratio was below average (approx. 0.96); thus, the "high" male range of digit ratio in this study may actually be average or below average compared to the normative population.)

Neave, Hamilton, and Fink (2007) reported the results of an analogous SNARC study to Bull and Benson's (i.e., greater lateralization) in females with low (masculine) digit ratios, and men with high (feminine) digit ratios on a test of subitizing. Their findings, in contrast to Bull and Benson's (2006), appear to fit the quadratic relationship model, in that females who were at the masculine end of the female distribution, and males who were at the feminine end of the male distribution were more functionally lateralized.

3.5.2 Sexually Dimorphic Brain Organization Pathways and Timescales?

An alternative way of explaining the 2D/4D data reviewed here is that males' and females' prenatal brain organization processes are affected differently by the same prenatal hormones. For instance, one study found the expected negative correlation between digit ratio and numerical competence in boys aged 6-11, but no such relationship for females (Fink, Brookes, Neave, Manning, & Geary, 2006). In another investigation, boys' digit ratios significantly predicted their absolute (Z numerical score) and relative (Z numerical – Z verbal) numerical scores on National Standard Assessment Tests in the United Kingdom. However, girls' digit ratios only predicted their absolute verbal abilities, with high digit ratio related to high

verbal ability. Similarly, digit ratio significantly predicted male, but not female performance in science classes (Romano et al., 2006). Meanwhile, prenatal testosterone measured via amniotic fluid from between the 14th and 20th weeks of gestation has been negatively correlated with girls', but not boys' later numerical competence (Finegan, Nichols, & Sitarenios, 1992). On the other hand, a recent study found an expected inverse correlation between digit ratio and spatial/numerical skills associated with an individual's undergraduate major in females, but not in males (Valla et al., 2010).

Although Fink et al. (2006) cite Finegan et al.'s study to support the claim that prenatal hormones differentially affect males and females, they question these findings due to the fact that prenatal testosterone levels in Finegan et al.'s study were from later in gestation than when 2D/4D is determined, thus raising the temporal asynchrony argument. Along these lines, it may even be the case that the lack of effects for females, but not males, in some of the above studies is due to different sensitive periods for prenatal brain organization in males and females. The curious findings in Finegan et al., then, may have been due to a female sensitive/critical period which is later in the gestational timetable than males'. Thus, a negative relationship between prenatal testosterone at approximately 16 weeks (the mean measurement time point in their study) and later numerical competence may indicate that prenatal testosterone is detrimental to females' numerical abilities, in contrast with the assumption that comparatively high prenatal testosterone (compared to other females' hormone levels) improves females' numerical abilities. Indeed, it would be consistent with these findings to argue that prenatal testosterone exposure may hinder any potential, presumably delateralized, alternative strategies females typically use for solving mathematical problems.

Questions of detrimental effects aside, there is additional support for sexually

differentiated brain organization critical periods that may explain the inconsistency in the predictive power of 2D/4D ratios in females' spatial/numerical abilities. For instance, in one study the spatial abilities (as measured by three tests: the Vandenberg Mental Rotations test, the Paper Folding test, and the Guilford-Zimmerman Spatial Orientation Test) of females who were classified as non-heterosexual were significantly higher than those of heterosexual females (Van Anders & Hampson, 2005). However, these same females' digit ratios were unrelated to their scores on all three spatial tests, and there was no difference between the digit ratios of the heterosexual and non-heterosexual females. Commenting on the developmental asynchrony of the 2D/4D ratio and spatial ability for both sexes, Puts, McDaniel, Jordan, and Breedlove (2008) suggest that sexual orientation may be more contemporaneous with spatial ability than the 2D/4D ratio. Even if true, however, this leaves a lacuna in the early brain organization position, which posits a positive manifold of correlations between all biomarkers by the onset of early adolescence.

3.5.3 Prenatally-Determined Ability, or Prenatally-Determined Preference?

There is also the possibility that the inconsistent relationship between prenatal testosterone (as measured by the 2D/4D ratio) and spatial and numerical competencies is due to an alternative pathway of influence of prenatal testosterone on spatial/numerical cognition. Namely, prenatal testosterone may affect preferences rather than directly establishing ability. Inconsistent findings, then, might be a function of the fact that, while ability and preference can go hand-in-hand, this is not always the case. For example, when presented with arrays of masculine and feminine objects and toys, men and women who showed more fixation on masculine objects had better targeting abilities and lower 2D/4D ratios (Alexander, 2006). In the same study, however, visual fixation times were unrelated to mental rotation ability, in both

sexes. In addition, there was no difference between individuals with higher masculine object fixations and those with higher feminine object fixations in retrospective reports of gender-linked childhood activities. Weis, Firker, and Hennig (2007) gave participants occupational interest inventories in addition to measuring their digit ratios. In males, there was a negative correlation between 2D/4D ratio and interest in "Realistic" and "Enterprising" careers, while females' digit ratios were negatively correlated with interests in "Enterprising" and "Investigative" careers. It would be hard to argue *a priori* that "Realistic" typifies math-intensive careers any more than "Investigative".

These empirical attempts to unravel digit ratio correlations with preference versus ability are consistent with the causal pathway hypothesized by Ceci et al., (2009) to exacerbate inherent, but initially very small, sex differences in spatial and numerical domains. Others have shown a "snowball effect" whereby initial cognitive ability breeds preference, which feeds back into enlarged ability differences (Dickens & Flynn, 2001), but it is equally easy to hypothesize that inherent preference spawns ability differences, before feeding back into still greater differences in preference.

At the same time, Alexander's (2006) findings imply a more powerful idea than this "snowball" causal model, because the finding that digit ratio was related to masculine/ feminine object preferences but not to mental rotation ability, coupled with the fact that no mental rotation ability differences existed between individuals preferring masculine versus feminine objects, suggests the possibility that preference and ability can be dissociated, even into adulthood. More importantly, it suggests that the influence of sex difference due to prenatal testosterone exposure is not directly on ability, but emerges as a function of interest. Given that no childhood activity differences existed between those preferring masculine versus feminine objects, this suggests

that children's play preferences may be (in cases of dissociation between interest and ability) functions of interest, and not ability. One possible explanation for this is that play activities are influenced by socialization, but socialization may not alter preferences, which may be rooted in prenatal testosterone exposure. Relatedly, behaviors such as risk-taking and vigilance that are associated with early testosterone and a lower male digit ratio are relevant in financial trading, where male ratios predicted profit and loss records of 49 traders in London (Coates, Gurnell, & Rustichini, 2009).

A recent example of the inherent ambiguity involved in distinguishing between results due to prenatal testosterone exposure versus subsequent activities is the study reported by Vuoksimaa et al. (2010) in which she and her colleagues concluded that "We...found that females with male co-twins outperformed females with female co-twins (on a mental rotation test). Our results are consistent with the prenatal masculinization hypothesis, according to which masculinization occurs in females with male co-twins as a result of intrauterine exposure to testosterone." Notwithstanding the authors' statement, it appeared to be the case that having a female co-twin enhanced the mental rotation ability of her male co-twin (it was not clear from their figure 1 if this effect was significant). If significant, one could just as readily claim that feminization is good for mental rotation ability, calling further into question the welter of expected correlations between prenatal hormones, sex differences in spatial ability, and the digit ratio. Or, alternatively, one could claim that having a male co-twin enhanced females' subsequent spatial ability because of the subsequent influence of being exposed to male activities (blocks, erector sets, Legos) that otherwise might be less evident if one's sib is another female. We are not advancing either of these claims; we merely invoke them as evidence of the ambiguity of some of the key findings supporting the organizational role of hormones in

developmental outcomes.

In any event, a hormonally-induced sex difference in preferences that is functionally independent of actual ability jibes with Ceci et al.'s assertion that female choice (e.g., to enter the humanities rather than science fields, to prioritize family over a career in science, or to enter a non-mathematical scientific instead of a math-intensive field) is likely much more important than biological differences in ability when explaining female underrepresentation in mathematically-intensive fields. This helps explain why, despite a 2-to-1 ratio of males to females scoring among the top 1% of the mathematics distribution, there is nowhere close to one-third women occupying math-intensive positions in fields such as physics, engineering, computer science, economics, chemistry, and mathematics. Clearly, more than high levels of math aptitude are at play in sorting men and women into careers.

3.6 Discussion

In this paper we have attempted to synthesize findings across many areas of research (endocrinology/psychoneuroendocrinology, neuroscience/anatomy, mental rotation, sociology, evolutionary psychology, education, personality, and genetics) comprising the long chain of logic underlying the use of the 2D/4D ratio to make claims about sex differences in brain organization. Such a synthesis is both important and timely, as the comparative ease and non-invasiveness with which 2D/4D can be measured has made it an increasingly central piece of evidence in the biologically-based cognitive sex differences and the Brain Organization Hypothesis. Yet the ease of 2D/4D measurement has revealed its weakness, and it is now readily used to support inconsistent conclusions about sex differences in brain organization without a full understanding of how tenuous the evidence supporting its validity actually is. Indeed, in carrying out this review we have identified numerous inconsistencies and have raised alternative

interpretations of the 2D/4D ratio studies that, taken together, temper claims related to early brain organizational theory. Although nothing in this review refutes this theory, far stronger support for it is needed if it is to account for the dearth of women in mathematical fields. In the remainder of this paper we identify specific areas and methodological issues in need of resolution, based on our review of the literature. (A summary of the reviewed studies can be found in Table 7).

3.6.1 Sample Representativeness

As noted in several places, some studies attempting to delineate high- and low digit ratios within-sex did so with unrepresentative samples. If the seemingly contradictory spatial advantage for high digit ratio (i.e. more feminized) males is actually an artifact of using a male sample that has a below-average ratio to begin with, then the ideal ratio may in this hypothetical example be the male average. Such potential confounds only confuse interpretations and claims (e.g., the super-male brain supposedly being more androgenized) and underscore the need for better sampling. More broadly, the issue of sample representativeness looms as a barrier to theory validation.

3.6.2 Reliability and Validity of 2D/4D in General, and as Proxy for Degree of Lateralization

Although we have reviewed the digit ratio-cognitive sex difference literature under the assumption that digit ratio is a valid and reliable proxy for prenatal androgen exposure, such assumptions may be premature. That is, there is some evidence that digit ratio is fluid, and fluctuates, for instance, with menstrual cycle (Mayhew, Gillam, McDonald, & Ebling, 2007). There are also larger within-sex but between-ethnicity 2D/4D differences in digit ratio than there are between-sex differences, raising questions about the magnitude of testosterone required to affect both physiological and psychological outcomes (Cohen-Bendahan, van de Beek, &

Berenbaum, 2005).

While 2D/4D is an alluringly simple tool for studying the effects of prenatal testosterone on development, hence accounting for the hundreds of studies employing it (Voracek & Loibl, 2009), researchers may have "put the wagon before the horse" in using this methodology. As noted earlier, it seems vastly over-interpreted considering that it is indicative – at best – of prenatal testosterone level at one particular gestation period when limb and urogenital functions are most affected. Considering that we currently do not yet fully know when cognitive, athletic, and behavioral attributes are determined *in utero*, digit ratio determination may only coincide with a fraction of traits determined prenatally. Puts et al. (2008) made a similar point in noting the asynchrony of digit growth and mental rotation ability. We currently lack an understanding of the gestational windows of these attributes, and thus risk arriving at conclusions about cognitive and behavioral developments that at times correlate with digit ratio. More validation is needed via longitudinal amniocentesis studies that are later compared with digit ratios and other attributes; this will provide a clearer picture of how, when, and what prenatal testosterone affects. In this regard, if such studies are conducted it will be important to address criticisms of Baron-Cohen's methodology (e.g., Jordan-Young, 2010).

3.6.3 Unitary Nature of Spatial and Numerical Abilities, Functionally and Methodologically

More care should also be taken in selecting appropriate and, if possible, novel tests of spatial and numerical abilities. Given findings that mental rotation performance may have less to do with the process of rotation than is assumed (Hooven, Chabris, Ellison, & Kosslyn, 2004), and the claim that spatial and numerical competencies can be reduced to the unitary dimension of spatially-represented magnitude (Bull & Benson, 2006), there may be more incisive method than those currently used to investigate exactly where any cognitive sex differences exist. If spatial

and numerical abilities (and, within spatial abilities, 2-D and 3-D mental rotation) are not as unitary as Bull and Benson assert, then using a wide array of different tests is necessary in future investigations, so that interpretations do not rest upon islets of ability.

3.7 Conclusion

In sum, as with the broader debate about Brain Organization Theory, the 2D/4D-spatial/numerical ability relationship is complicated. Evidence has been provided for three alternative possibilities to the currently assumed relationship, all of which can be reasonably inferred from the extant 2D/4D-cognitive sex difference literature. First is a quadratic, inverse-U relationship, purported by studies such as Brosnan (2006). Unfortunately, contradictory evidence exists not only between studies but even within studies such as Brosnan's, as discussed above.

Second is the possibility of sexually-differentiated prenatal brain organization pathways and timescales, which may manifest in sex-dependent relationships between, for instance, digit ratio and spatial/numerical ability (e.g., Valla et al., 2010). To date, research has been limited to addressing the existence of this relationship, based on the questionable assumption that there are no sex differences in these prenatal developments and brain organization pathways; alternative hypotheses including different developmental trajectories for males and females should be entertained.

Third is a possible relationship between prenatal hormone exposure and preferences, rather than ability. Future studies should attempt to differentiate between biologically-influenced interest, and biologically-influenced ability, which may be dissociable. Separating the effects of interests and abilities may help determine what, if any, inherent sex differences in spatial and numerical ability actually exist, and may be a promising area of reconciliation between the nature and nurture sides of this debate.

Support for early brain organization theory will require closer interleaving of correlational and experimental evidence. Most of the support for this theory has come from studies that were designed in response to a handful of prior studies – which is how good science accumulates. However, in the case of early brain organization theory, studies are needed that respond to the topography of the theory rather than to one of its high-relief landmarks. Future research will need to be framed in terms of the larger theory instead of local hypotheses. Such an approach will protect against short-sightedness that sometimes characterizes this literature.

3.8 Relevance to Work Three

As discussed in the first work, autism is in many ways a "natural experiment" in cognitive and social variation, an extreme version of the hypothesized inverse relationship between systemizing and empathizing skills, as well as an example of the cognitive and social effects of extreme levels of prenatal androgen exposure. Whereas the two works presented so far have concerned normative variation and co-variation in empathizing and systemizing, and their neurobiological underpinnings, the third work re-contextualizes these aspects of the E-S Co-variance Hypothesis as normative extensions of longstanding ideas in autism research.

Specifically, the so-called "Kanner Hypothesis" of autism holds that seemingly unrelated autistic social and non-social traits – some of which can be characterized as high systemizing or low empathizing – share a common underlying cause. The importance of the first two works with regards to the Kanner Hypothesis, then, lies in the ideas that 1) E-S Co-variance is greater in males than in females; and 2) Sex-dependent E-S Co-variance has prenatal neurobiological origins, but these prenatal effects may have their greatest influence indirectly, through initiating sexually dimorphic developmental trajectories. Together these ideas suggest that E-S Co-variance increases with greater prenatal androgen exposure, but also that much of E-S Co-

variance arises as a function of development.

The third work takes the evidence of E-S Co-variance from typical populations discussed in the preceding works and brings it together with evidence of E-S Co-variance from autistic symptom factor structures found in clinical investigations. The Interactive Specialization model of neurological and behavioral development that results bridges the gap between E-S Co-variance patterns found in typical and autistic populations.

REFERENCES

Alexander, G. (2006). Associations among gender-linked toy preferences, spatial ability, and digit ratio: evidence from eye-tracking analysis. *Archives of Sexual Behavior, 35,* 699-709.

Baron-Cohen, S. (2003). *The essential difference: The truth about the male and female brain.* London: Penguin.

Berenbaum, S., & Resnick, S. (2007). The seeds of career choices: Prenatal sex hormone effects on psychological sex differences. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 147–158). Washington, DC: American Psychological Association.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature–nurture in developmental perspective: A bioecological theory. *Psychological Review, 101,* 568–586.

Brosnan, M. (2008). Digit ratio as an indicator of numeracy relative to literacy in 7 year old British schoolchildren. *British Journal of Psychology, 99,* 75-85.

Brosnan, M. (2006). Digit ratio and academia: Implications for the relationship between prenatal testosterone and academia. *British Journal of Psychology, 97,* 455-466.

Brookes, H., Neave, N., Hamilton, C., & Fink, B. (2007) Digit ratio (2D/4D) and lateralization for basic numerical quantification. *Journal of Individual Differences, 28,* 55-59.

Bull, R., & Benson, P. (2006). Digit ratio (2D/4D) and the spatial representation of magnitude. *Hormones and Behavior, 50,* 194–199.

Buss, D., & Schmitt, D. (1993). Sexual strategies theory: an evolutionary perspective on human mating. *Psychological Review, 100,* 104–32.

Ceci, S. J. (1996). *On intelligence: A bioecological treatise*. Cambridge, MA: Harvard University Press.

Ceci, S. J. & Williams, W. M. (2010a). *The Mathematics of Sex: How biology and society conspire to limit talented women and girls.* NY: Oxford University Press.

Ceci, S. J. & Williams, W. M. (2010b). Sex Differences in Math-Intensive Fields. *Current Directions in Psychological Science, 19(5)*, 275-279.

Ceci, S. J. & Williams, W. M. (in press). Understanding Current Causes of Women's Underrepresentation in Science. *Proceedings of the National Academy of Sciences*.

Ceci, S. J., Williams, W.M., & Barnett, S.M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135,* 218-261.

Cohen-Bendahan, C.C., van de Beek, C., & Berenbaum, S.A. (2005). Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. *Neuroscience and Biobehavioral Reviews, 29,* 353–384.

Coates, J.M., M. Gurnell, M., & Rustichini, A. (2009). Second-to-fourth digit ratio predicts success among high-frequency financial traders. *Proceedings of the National Academy of Sciences, 106,* 623-628

Dickens, W.T., Flynn, J.R., (2001). Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychological Review 108*, 346–369.

Else-Quest, N., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103–127.

Falter, C. M., Arroyo, M., & Davis, G. J. (2006). Testosterone: activation or organization of spatial cognition? *Biological Psychology, 73,* 132–140.

Feingold, A. (1994). Gender differences in variability in intellectual abilities: A cross-cultural perspective. *Sex Roles, 30,* 81–92.

Finegan, J., Niccols, G. A., & Sitarenios, G. (1992). Relations between prenatal testosterone

levels and cognitive abilities at 4 years. *Developmental Psychology, 28,* 1075-1089.

Fink, B., Brookes, H., Neave, N., Manning, J.T., & Geary, D.C. (2006). Second to fourth digit ratio and numerical competence in children. *Brain and Cognition, 61,* 211-218.

Gallagher, A. M., & Kaufman, J. C. (2005). *Gender differences in mathematics: An integrative psychological approach.* Cambridge, United Kingdom: Cambridge University Press.

Geschwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization: Biological mechanisms, associations, and pathology.* Cambridge, MA:MIT Press.

Guo, G., & Stearns, E. (2002). The social influences on the realization of genetic potential for intellectual development. *Social Forces, 80,* 881–910.

Halpern, D. F., Benbow, C. P., Geary, D., Gur, R., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8,* 1–51.

Halpern, D.F. (2004). A cognitive taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13*, 135–139.

Hönekopp, J., Bartholdt, L., Beier, L., & Liebert, A. (2007). Second to fourth digit length ratio (2D/4D) and adult sex hormone levels: New data and a meta-analytic review. *Psychoneuroendocrinology, 32,* 313–321.

Hooven, C., Chabris, C., Ellison, P., & Kosslyn, S. (2004). The relationship of male testosterone to components of mental rotation. *Neuropsychology, 42,* 782–790.

Hyde, J. S. (2005). The gender similarity hypothesis. *American Psychologist, 60,* 581–592.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321,* 494–495.

Hyde, J. S. & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings*

of the National Academy of Science, 106, 8801-8807.

Jordan-Young, R. (2010). *Brain storm: The Flaws in the Science of Sex Differences*. Cambridge, MA: Harvard University Press.

Kempel P, Gohlke B, Klempau J, Zinsberger P, Reuter M, & Hennig J. (2005). Second-to-fourth digit length, testosterone and spatial ability. *Intelligence, 33,* 215-230.

Kondo, T., Zákány, J., Innis, J., & Duboule, D. (1997). Of fingers, toes and penises. *Nature, 390,* 29-30.

Lohman, D., & Lakin, J. (2009). Consistencies in sex differences on the cognitive abilities test across countries, grades, test forms, and cohorts. *British Journal of Educational Psychology, 79,* 389-407.

Lutchmaya, S., Baron-Cohen, S. & Raggett, P. (2002). Foetal testosterone and eye contact in 12-month old infants. *Infant Behavior and Development, 25,* 327–335.

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

Manning, J. (2002). Digit ratio: A pointer to fertility, behavior and health. New Brunswick, NJ: Rutgers University Press.

Manning, J. and Leinster, S. (2001). 2nd to 4th digit ratio and age at presentation of breast cancer. *The Breast*, *10*, 355–35.

Manning, J., Barley, L., Walton, J., Lewis-Jones, D., Trivers R., Singh D, et al. (2000). The 2^{nd} and 4th digit ratio, sexual dimorphism, population differences, and reproductive success: evidence for sexually antagonistic genes? *Evolution & Human Behavior, 21*,163– 83.

Manning, J., Trivers, R., Thornhill, R., & Singh, D. (2000). The 2nd:4th digit ratio and asymmetry of hand performance in Jamaican children. *Laterality, 5,* 121-132.

Mayhew, T. M. L., Gillam, R., McDonald, J., & Ebling, F. J. P. (2007). Human 2D (index) and 4D (ring) digit lengths: Their variation and relationship during the menstrual cycle. *Journal of Anatomy, 211*, 630-638.

Moffat, S. & Hampson, E. (1996). A curvilinear relationship between testosterone cognition in humans. *Psychoneuroendocrinology 21,* 323 – 337.

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

Mullis, I. V. S., Martin, M. O., Fierros, E. G., Goldberg, A. L., & Stemler, S. E. (2000, July). Gender differences in achievement: IEA's third international mathematics and science study (TIMSS). Chestnut Hill, MA: Boston College, Lynch School of Education, International Study Center. Retrieved March 18, 2008, from http://timss.bc.edu/timss1995i/TIMSSPDF/t95_gender_all.pdf *OECD. (2004). Learning for tomorrow's world: First results from PISA 2003.*. OECD: Paris.

Nelson D, Brammer C (2010) *A national analysis of minorities in science and engineering faculties at research universities*. Retrieved June 17, 2010, from http://chem.ou.edu/~djn/diversity/faculty_Tables_FY07/FinalReport07.html

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., & Bar-Anan, Y. et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Science, 106*, 10593-97.

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,* 47-54.

Penner, A. M. (2008). F *American Journal of Sociology, 114,* S138–S170.

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

Romano, M., Leoni, B., & Saino, N. (2006). Examination marks of male university students positively correlate with finger length ratios (2D/4D). *Biological Psychology, 7,* 175-182.

Shalala, D. E., Agogino, A. M., Bailyn, L., Birgeneau, R., Cauce, A. M. Deangeles, C., et al. (2007). *Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering.* Washington, DC: National Academies Press.

Shonkoff JP, Phillips D, et al; National Research Council. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development.* Washington, DC:National Academy Press.

Sommers, C. H. (2008, March/April). Why can't a woman be more like a man? *The American, 2*. Retrieved September 19, 2008, from www.american.com/ archive/2008/march-april magazine contents/why-can2019t-a-woman-be-more-like-a-man

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist, 60*, 950–958.

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in Cognitive Abilities Test scores: A UK national picture. *British Journal of Educational Psychology, 76*, 463–480.

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

Valla, J.M., Ganzel, B.L., Yoder, K.J., Chen, G.M., Lyman, L.T., Sidari, A.P. Keller, A.E., Maendel, J.W., Perlman, J.E., Wong, S.K.L., & Belmonte, M.K. (2010). More than maths and mindreading: sex differences in empathizing-systemizing covariance. *Autism Research, 3,* 174-184.

Voracek, M. & Loibl, L. (2009). Scientometric analysis and bibliography of digit ratio (2D/4D) research, 1998-2008. *Psychological Reports, 104*, 922-956.

Vuoksimaa, E., Kaprio, J., 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*, 1069-1071.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30-year examination. *Intelligence, 38*, 412-423.

Wainer, H. & Steinberg, L. (1992). Sex differences in performance on the mathematics section of the Scholastic Aptitude Test: A bidirectional validity study. *Harvard Educational Review, 62*, 323-326.

Weis, S. E., Firker, A., & Hennig, J. (2007). Associations between the second to fourth digit ratio and career interests. *Personality and Individual Differences, 43,* 485-493.

Wisniewski, A. (1998). Sexually-dimorphic patterns of cortical asymmetry, and the role for sex steroid hormones in determining cortical patterns of lateralization. *Psychoneuroendocrinology, 23,* 519–547.

Witelson, S., & Nowakowski, R. (1991). Left out axons make men right: A hypothesis for the origin of handedness and functional asymmetry. *Neuropsychologia, 29,* 327–333.

CHAPTER 4

CO-VARIATION OF SOCIAL AND NON-SOCIAL 'AUTISTIC' TRAITS, IN AUTISTIC AND TYPICAL DEVELOPMENT: A REVIEW AND SYNTHESIS

4.1 Introduction

Current consensus holds that autism represents the extreme of a spectrum subsuming certain cognitive, social, and behavioral characteristics. These characteristics are categorized behaviorally in terms of the diagnostic "triad" of social interaction deficits, communicative deficits, and restricted and repetitive behaviors and interests (American Psychiatric Association, 1994). They are described theoretically in terms of (1) deficits in Theory of Mind (ToM) or understanding intentionality (Baron-Cohen, 1995; Baron-Cohen, Leslie, & Frith, 1985) and in future-oriented cognitive flexibility, or executive functioning (EF) (Ozonoff, Pennington, & Rogers, 1991); and also (2) a bias toward local over global information processing, or weak central coherence (WCC) (Frith & Happé, 1994). In the terms of normative psychology, these two groups of traits are equivalent to Trope's notions of decreased "psychological distance" and low "level of construal," respectively (see Trope & Liberman, 2010). Less extreme positions along the autism spectrum are occupied by Asperger syndrome (AS), pervasive developmental disorder not otherwise specified (PDD-NOS), and the more subtle, sub-clinical idiosyncrasies of many first-degree relatives of individuals with autism, the sort described by Kanner in his seminal work (1943). This latter, broader autism phenotype has been empirically distinguished to such a degree that elements of it are sometimes referred to as the proper-noun Broader Autism Phenotype, or BAP (Baron-Cohen & Hammer, 1997a; Dawson et al., 2002; Happé et al., 2001; Lord, Cook, Leventhal, & Amaral, 2000; Piven, 1999).

4.1.1 Questioning the "Kanner Hypothesis"

Researchers and clinicians have long questioned the assumption that social (e.g., ToM deficits) and non-social (e.g., WCC) ASC traits are related, co-varying between individuals in a manner implying a shared underlying cause. As theorists have previously pointed out (Mandy & Skuse, 2008; Happé & Ronald, 2008), this presumption of social/non-social co-variance in ASC individuals dates to Kanner's original descriptions of autism (Kanner, 1943) that framed autistic social and detail-oriented cognitive bias as two sides of the same coin. For instance, in Kanner's initial synthesis of commonalities among his first eleven case studies, the children in question are said to share a tendency to "establish and maintain an excellent, purposeful, and 'intelligent' relation to objects that do not threaten to interfere with their aloneness, but are from the start anxiously and tensely impervious to people…," and "If dealing with another person becomes inevitable, then a temporary relationship is formed with the person's hand or foot as a definitely detached object, but not with the person himself. All of the children's activities and utterances are governed rigidly and consistently by the powerful desire for aloneness and sameness." Since that time, the 'Kanner hypothesis' - that, in Kanner's parlance, a preference for "aloneness" is the flipside of a related need for "sameness" - has persisted as one of the most hotly debated questions in autism research, being directly relevant to whether there is in any sense a single *cause* or, if sameness and aloneness are *not* related, *causes*, underlying the condition. This question of the unity of social aloneness and perceptual sameness was the implicit basis of the debate between ToM and WCC accounts of autism, as WCC proponents held that the ToM account did not adequately explain the non-social, perceptual aspects of ASC (Frith & Happé, 1994).

Kanner's definition of autism was, however, based on observations of patients presenting with both social and non-social symptoms to begin with, to the selective exclusion of cases

where social or non-social deficits were absent. Assessing the validity of theories of autism according to how well they account for social and non-social symptoms simultaneously is thus ultimately an exercise in circular logic, as Kanner's observations were based on patients within whom, by clinical diagnostic definition, social and non-social deficits are comorbid (Happé & Ronald, 2008; Happé, Ronald, & Plomin, 2006). As such, there is not necessarily reason to assume, as Kanner might, that severe social interaction deficits come part and parcel with extremity in repetitive behaviors and/or restricted interests. Similar criticisms of circular logic have been levied against studies (*e.g.,* Van Lang et al., 2006; Boomsma et al., 2008) basing Autism Diagnostic Interview-Revised (ADI-R) factor analyses solely on items tapping the DSM-IV diagnostic triad (Kamp-Becker, Ghahreman, Smidt, & Remschmidt, 2009). (For a historical review of the repeated re-emergence of Kanner's assumptions in autism research and theory, see Happé & Ronald, 2008).

4.1.2 The Kanner Hypothesis in the Normative Population

A separate line of research has extended the autism spectrum into and throughout the normative population, such that social and non-social autistic traits are normally distributed across all individuals, with social deficits and cognitive idiosyncrasies taking subclinical forms. On these dimensional axes, clinically diagnosable autism differs from normative functioning quantitatively and by degree, rather than being qualitatively distinct from normative cognitive variation. The best known of these descriptions is Empathizing-Systemizing (E-S) theory, which holds that ASC is an extreme version of the human brain in terms of its neuroarchitectural "male-ness," the prototypical "male brain" being organized for superior "systemizing" (Baron-Cohen, 2002; Baron-Cohen & Hammer, 1997b), or processes of observation-based rule-making reducing the world to a series of "lawful, finite, and deterministic" rules. The prototypical "female brain,"

by contrast, is better organized for "empathizing," the ability to accurately attribute and affectively respond to others' mental states (Baron-Cohen, 2002).

The full range of these attributes, from high- to low empathizing and high- to low systemizing, constitutes the entire E-S spectrum, within which the general population is normally distributed (Baron-Cohen & Hammer, 1997b). One's position on this continuum is thought to be determined partly by levels of prenatal testosterone exposure (Baron-Cohen, Lutchmaya, & Knickmeyer, 2004), with higher levels of prenatal exposure producing a more lateralized, systemizing-oriented, "male" brain (Witelson & Nowakowski, 1991; Geschwind & Galaburda, 1987), gestational overexposure to testosterone being only one amongst many interacting factors that can bias brain and cognitive development towards an autistic outcome. Supporting this assertion is evidence of higher testosterone levels in amniocenteses of individuals presenting with autism-like empathizing difficulties and systemizing prowess later in development (Baron-Cohen et al., 2004), and lower than average second-to-fourth digit length ratios (a biomarker proxy measure of prenatal testosterone exposure) among ASC individuals and their first degree relatives, individuals falling into the BAP category, and individuals with systemizing prowess and sub-clinical empathizing difficulties (see Valla & Ceci, 2011 for a full review).

As E-S theory was built upon – or, one might even say, built to account for – the assumptions of social and non-social co-variance in clinical populations that began with Kanner, it is worth questioning the Kanner hypothesis not only in the context of clinical presentation, but in typical development, too (Carroll & Chiew, 2006; Valla et al., 2010). If the E-S factor space comprises a single axis with high empathizing and low systemizing at one end and low empathizing and high systemizing at the other, as initially described (Baron-Cohen & Hammer, 1997b), then systemizing ability should carry with it *tradeoffs* in empathizing, and *vice versa*.

Such would be, in effect, a normative-variation extension of the Kanner hypothesis. An alternative model, implicit in subsequent work (Baron-Cohen, 2002), presents the E-S factor space as bidimensional, wherein empathizing and systemizing vary *independentl*y, an autistic, extreme male brain arising from concurrently low empathizing and high systemizing abilities, an extreme female brain arising from the opposite pattern. This alternative, two-dimensional model of the E-S factor space implies no tradeoffs between systemizing and empathizing.

The importance of E-S co-variance in understanding normative cognition cannot be understated. For one, an inverse E-S co-variance pattern in the normative population (i.e. the single-axis E-S model) would mean that mathematical and spatial skills or deficits would be predictive of such seemingly unrelated social deficits or skills as understanding intentionality. In lay terms, testing for this type of E-S co-variance in the normative population means questioning the stereotype of spatial and/or mathematical prowess going hand-in-hand with social ineptitude.

Insofar as any normative E-S co-variance patterns may hold also at the clinical extremity, evidence for or against E-S co-variance in the normative population would also be essential to consider alongside evidence for or against the Kanner hypothesis in clinical presentation. More to the point, E-S independence in typically developing individuals would bring into question theories attempting to account for ASC with a single-cause explanation, and the broader idea of ASC arising from coincidence of partially independent but synergistic factors might begin to seem more likely (Happé, Ronald, & Plomin, 2006). Thus, whilst understanding E-S co-variance patterns can tell us important things about normative functioning, it can also inform us on the causes underlying ASC.

Whilst the Kanner debate is framed here as one of evidence for *or* against E-S co-variance and the Kanner hypothesis, the distinction between co-variance on the one hand and

independence between autism's social and non-social symptom domains on the other need not be binary: social and non-social symptoms may be partially dependent, synergizing and reinforcing each other to some extent, but varying independently enough to give rise to a multitude of individual phenotypes. This independence may be particularly pronounced in cases or populations where levels of autistic traits are subtle and where, therefore, mutually reinforcing interactions amongst these symptom domains are weak. Thus independence of autism's various social and non-social symptom domains may be strongest not within but beyond the autism spectrum, in individuals with the Broader Autism Phenotype (Piven, 1999) and in the subtle levels of autistic traits that occur throughout the general population, and – to apply Baron-Cohen's terms – may be stronger in "female brain" individuals rather than the "male brain" which, according to Baron-Cohen, autism more closely approaches (Baron-Cohen et al., 2004). Evidence regarding E-S co-variance in normative samples is in this way at least as important as clinical sample data in assessing the veracity of the Kanner hypothesis.

Lastly, the use of an E-S theory framework is not meant to imply that E-S theory, its constructs, or its predictions are undisputed. On the contrary, a substantial portion of the data reviewed below brings into question central assumptions of E-S theory. In fact, challenging E-S theory is in a sense the purpose of this review, though the interest here lies more narrowly in challenging the Kanner assumption of co-variance between the "E" and the "S," rather than evaluating E-S theory in general. The assumption of Kanner co-variance extends well beyond E-S theory, and the underlying questions are the same whether the discussion is framed in terms of 'Empathizing' versus 'Systemizing', or iconic, detail-oriented, egocentric perception, cognition and action versus symbolic, contextualized, allocentric perception, cognition and action. The present discussion is framed in E-S terms because E-S theory seems the most fully articulated

example of Kanner co-variance assumptions, within and beyond the spectrum, in contemporary autism theory and research. Indeed, the present review is not the first to recognize the utility of an E-S framework in this context (Mandy & Skuse, 2008).

The use of E-S terminology nonetheless necessitates two key qualifications. First, the terms 'male brain' and 'female brain' are perhaps unfortunate in that a complete identification of brain type with sex and/or gender can mislead; the initial elaboration of E-S theory came with the caveat that many females would be classified as having male-type brains, and many males as having female-type brains (Baron-Cohen et al., 2004). Second, in terms of the idea of empathizing 'deficits' in autism, the term 'empathizing' in this case refers specifically to *cognitive* empathy, rather than *affective* empathy; the difference between the two is the difference between explicit perspective taking and emotional contagion.

4.1.3 Relevant Population Characteristics and Evidence for Assessing the Kanner Hypothesis

The normal distribution of autistic traits throughout the clinical and non-clinical population (Baron-Cohen & Hammer, 1997b) necessitates that factor structures and co-variance patterns in clinical, subclinical, and non-clinical samples be considered alongside one another. The developmental nature of autism necessitates inclusion of evidence from across age groups, as social/non-social co-variance patterns present early in development (i.e. close to the prenatal period) may differ greatly from patterns emerging later, after any comorbid social and non-social traits repeatedly interact behaviorally and cognitively.

Measures most directly relevant to assessing the Kanner hypothesis are trait inventories, and cognitive and social psychometrics. Trait inventories include clinically-oriented symptom inventories (e.g., Autism Diagnostic Inventory - Revised, or ADI-R); behavioral inventories applicable to clinical, subclinical, and non-clinical populations (*e.g.,* Social Responsiveness

Scale, or SRS); and self-reported assessments of sub-clinical manifestations of autistic symptoms, in the form of personality traits, behaviors, preferences, and tendencies used mainly with sub-clinical and normative populations [e.g., Autism Spectrum Quotient, or AQ (Baron-Cohen et al., 2001); the Systemizing Quotient, or SQ (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003); and the Empathizing Quotient, or EQ (Baron-Cohen & Wheelwright, 2004)]. Interpretations of ASC trait structure based on the AQ, EQ, and SQ come with the caveat that these measures do not capture all ASC traits, being based on the higher functioning end of the autism spectrum (e.g., Asperger Syndrome, BAP). The AQ, for instance, contains few high functioning analogues of the sensory abnormalities and motor stereotypies seen in clinical cases (Valla et al., 2010).

Relevant cognitive measures include those roughly categorized along systemizing and empathizing lines. In the systemizing category are assessments of spatial skills and pattern recognition (e.g., Block Design and Matrix Reasoning subtests of the Wechsler Intelligence Scale for Children, or WISC-III) and other measures favoring cognitive inflexibilities and detail-oriented biases, such as disembedding speed on the Embedded Figures Test (EFT; Witkin, 1950), and susceptibility to visual illusions (*e.g.,* Walter, Dassonville, & Bochler, 2008). In the empathizing category are assessments of facial mental state/emotion reading (Reading the Mind in the Eyes Test, or RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), face recognition (*e.g.,* Benton face recognition test; Benton, Hamsher, Varney, & Spreen, 1983), and ToM (*e.g.,* false belief test; Wimmer & Perner, 1983).

Relevant analytic methods can be roughly separated into two complementary approaches: factor analyses of the underlying trait structure of social and non-social ASC-related traits *within* individuals, and correlational analyses assessing co-variance of these traits *between* individuals.

The latter approach allows social/non-social relationships to be tested more directly than do factor analyses, wherein social/non-social relationships are inferred from factor loadings. Factor analytic methods used in the studies reviewed below can be further divided into exploratory (EFA) and confirmatory (CFA) subtypes. CFA is a top-down, theory driven, model-fitting method that tests the accuracy of imposed factor structures, in terms of their ability to account for individual variation in a given behavioral inventory. EFA, meanwhile, exposes relationships between ASC traits in a data-driven, bottom-up manner which complements correlational studies and CFAs by making no presuppositions about factor structure. Factor analytic studies in the literature use EFA, CFA and, in some cases, EFA validated by CFA using a second, independent sample.

4.2 Review of Evidence for and Against the Kanner Hypothesis

A review of the literature finds significant evidence against the Kanner hypothesis, though the evidence in favor is not insubstantial. Three interrelated themes emerge, which together reconcile these two seemingly opposing sets of evidence for and against the Kanner hypothesis, suggesting a model in which initially independent social and non-social traits interact during developmental of behavior and cognition, giving rise to emergent Kanner-like co-variance in behavioral presentation.

The three themes are as such: First, evidence suggests an alternative triad of primary, independent ASC factors – *Social Interaction Deficits, Cognitive Inflexibility*, *and Sensory Abnormalities* - which may more accurately reflect the underlying structure of autistic traits than the currently applied diagnostic triad. The proposed alternative triad better accounts for heterogeneities within two domains of the diagnostic triad – *Communicative Deficits* and *Restricted and Repetitive Behaviors and Interests* – heterogeneities that can produce spurious

appearances of Kanner-like social/non-social co-variance in some studies. Second, these three alternative-triad factors may be independently heritable, implying multi-factor causation at the biological level. Third, despite this biologically based independence which may seem to negate the Kanner assumption, Kanner-like social/non-social co-variance in *behavioral* presentation within individuals can arise and amplify during postnatal development. Three additional patterns in the data support this model: increased Kanner-like co-variance with increases in trait extremity in multiple domains; greater Kanner-like co-variance in males compared to females; and continuity in this pattern of increasing behavioral co-variance across non-clinical, subclinical, and clinical populations.

In the dialectic that follows, we begin by reviewing the evidence in terms of broad categories of findings for and against the Kanner hypothesis. The reviewed evidence is then synthesized, first into the three aforementioned themes, then into an overarching developmental model accounting for these themes. The explanatory power of the hypothesized developmental model then is demonstrated, first by extrapolating to an in-depth, novel account of autistic behavioral stereotypies; and second, in a more brief and speculative fashion, with potential developmental accounts of other traits – such as musical savantism and communicative deficits – that have tended in the past to encourage single-cause accounts of autism.

4.2.1 Evidence Against the Kanner Hypothesis

How does translation from the diagnostic triad to the proposed alternative triad transform evidence appearing to support the Kanner hypothesis into evidence against Kanner-like social/non-social co-variance?

Empirically, the *Communicative Deficits* factor of the diagnostic triad seems to consolidate incorrectly two separable types of deficits: those related to social reciprocation *per*

se, and communicative-domain instances of more domain-general cognitive inflexibilities (*e.g.,* verbal stereotypies). Factor analyses of several trait inventories converge on this fractionation between symptoms subsumed under the *Communicative Deficits* category. In a principal components analysis of the ADI-R (Georgiades et al., 2007), delays in language development loaded on a *Social Communication* factor, whereas repetitive and idiosyncratic speech loaded on an *Inflexible Language and Behavior* factor along with non-communicative, cognitive inflexibility-related behaviors (*e.g.,* circumscribed interests, ritualistic behaviors). Although Tadevosyan-Leyfer et al. (2003) found a more granular factor structure comprising *Spoken Language*, *Social Intent*, *Sensory Aversions, Savant Skills,* and *Compulsions* in their EFA of the ADI-R, inter-factor correlations revealed the *Spoken Language* factor was more highly related to both *Social Intent* (0.45) and *Compulsions* (0.33) factors than the other factors were related to each other (Table 1), singling out *Spoken Language* as a partial combination of a social and a domain-general factor, and suggesting that the appearance of Spoken Language as a distinct factor might have more to do with the structure of the test instrument and the clinical diagnostic definition of the syndrome than with the underlying trait structure. Two additional factor analyses of the ADI-R each resulted in two-factor models aligning social reciprocation aspects of *Communication Deficits* with other social deficits, and cognitive inflexibility aspects with a *Restricted and Repetitive Behaviors and Interests* factor (Snow, Lecavalier, & Houts, 2009; Frazier, Youngstrom, Kubu, Sinclair, & Rezai, 2008). These analyses were notable in terms of their large sample sizes ($N = 1861$, $N = 1170$, respectively).

Even though a factor analysis of the CARS revealed a distinct *Social Communication* factor, as was the case with the aforementioned ADI-R analysis this factor was highly correlated with both *Social Interaction* (0.529) and the non-social *Stereotypies and Sensory Abnormalities*

(0.415) factors (Magyer & Pandolfi, 2009). Similarly, among Child Asperger Syndrome Test (CAST) diagnostic triad subscores, *Communication Deficits* correlated more highly with *Social Interaction Deficits* and with *Restricted and Repetitive Behaviors and Interests* (0.39, 0.42, respectively, in males; 0.26, 0.34 in females) both, than *Social Interaction Deficits* were correlated with *Restricted and Repetitive Behaviors and Interests* (0.27 in males, 0.17 in females) (Ronald et al., 2006b). Such correlational differences would be expected from the mixture of social-communicative and cognitive-inflexibility symptoms lumped into the *Communication Deficits* factor of the current diagnostic triad.

In two independent factor analyses of items comprising the *Restricted and Repetitive Behaviors and Interests* subscale of the ADI-R, an *Insistence on Sameness* component was significantly correlated with the *Communication Deficits* domain subscore of the ADI-R, but not the *Social Interaction Deficits* domain subscore (Szatmari et al., 2006; Lam et al., 2008). Another ADI-R analysis (Van Lang et al., 2006; cross-culturally validated in Boomsma et al., 2008) found that a triad of *Impaired Social Communication*, *Stereotyped Features in Behavior and Communication*, and *Impaired Play Skills* outperformed both the DSM-IV triad and a single-factor model. Although the third factor of our hypothesized triad is *Sensory Abnormalities* where in this investigation it was *Impaired Play Skills* (e.g., difficulties with social play and forming peer relationships), the *Communicative Deficits* factor in both cases is split between social reciprocity difficulties and cognitive inflexibility.

A retrospective-versus-current presentation longitudinal analysis of child ADI-R scores found that both verbal and non-verbal repetitive behaviors fractionated into two distinct factors depending on level of complexity: stereotyped utterances and echolalia loaded on a *Stereotypies* factor along with motor stereotypies, whereas verbal rituals and neologisms loaded on a domain-

general cognitive inflexibility factor *Anxiety and Compulsions* along with restricted interests. Communicative reciprocity (*i.e.* give-and-take conversational skills), on the other hand, loaded on a *Social-Communication* factor (Kamp-Becker, Ghahreman, Smidt, & Remschmidt, 2009). On ADOS-G inventories of current (6-24y) symptom presentation, similar loadings mirrored the same distinctions between deficits of social reciprocation, and cognitive inflexibility: language ability loaded negatively on an *Interests and Compulsions* factor, and speech abnormalities and stereotyped behaviors loaded on a *Non/Verbal Behavior* factor, whereas conversational communication skills loaded on a *Social-Communication* factor. This combination of results at different ages indicates some developmental continuity of heterogeneity in *Communicative Deficit* domain symptoms.

Matson and colleagues' (2010) factor analysis of Baby and Infant Screen for Children with Autism Traits (BISCUIT)-Part 1, an inventory for ~1-3y infants and children, reveals an even more telling developmental picture. Although an EFA revealed a three-factor solution similar to the diagnostic triad, the vast majority of items loaded on *Socialization/Non-verbal Communication* and *Repetitive Behavior/Restricted Interests* factors, with multiple reciprocation-related communicative behaviors loading on the former factor, and inflexibility-related communicative deficits loading on the latter factor. This fractionation further supports heterogeneity among *Communicative Deficit* behaviors, but also highlights the utility of a developmental lens in understanding how and when a heterogeneous *Communicative Deficit* factor emerges.

Cognitive inflexibilities subsumed under the *Communicative Deficits* diagnostic domain may be fed by deeper roots. Repetitive verbal behaviors (*e.g.,* verbal rituals) and verbal behaviors unrelated to cognitive flexibility (*e.g.,* presence of phrase speech) both are less

variable within versus between affected sibships, whereas no such significant sibship effect exists for *Communicative Deficits* domain scores amalgamating flexibility-related and unrelated communicative skills (Silverman et al., 2002). This statistical dissociation has been interpreted as implying that these two types of "communicative" (or at least verbal) symptoms arise from independent genetic origins (Silverman et al., 2002). In a factor analysis of a new symptom inventory assessing autism-relevant endophenotypes, verbal stereotypies (*e.g.,* echolalia) loaded on a *Stereotypic Behavior* factor along with non-verbal, motor stereotypies (Sacco et al., 2010). The same study's loading of "age at first social smile" on the cognitive inflexibility-related *Stereotypic Behavior* factor seems, though, to violate a factorization of social reciprocity versus cognitive flexibility-related communicative deficits; this relationship is, one could argue, the Kanner hypothesis in its most basic form. This seemingly contradictory finding is re-interpreted in more detail below.

Assuming factorial continuity between clinical and non-clinical populations, further support for heterogeneity within the *Communicative Deficits* domain comes from studies on typically developing populations using the AQ (Austin, 2005): Of the "Big Five" personality traits (Openness to new experience; Conscientiousness; Extraversion; Agreeableness; and Neuroticism, well-established dimensions of personality), AQ Communication subscores are significantly positively correlated with Neuroticism, and negatively correlated with Extraversion, Openness, Agreeableness, and Conscientiousness. Meanwhile, AQ Social Skills subscores are related (negatively) only to Extraversion and Agreeableness, and AQ Attention to Detail subscores are related (negatively) only to Openness and Conscientiousness. AQ Communication subscores, in other words, describe personality constructs ostensibly related to cognitive (in)flexibility (Neuroticism, Conscientiousness, and Openness) as well as traits related to

willingness to engage in social interaction (Extraversion and Agreeableness). AQ Social Skills and Attention to Detail subscores, meanwhile, do not overlap in their Big Five trait loadings, yet both overlap with AQ Communication subscores (Austin, 2005). Even though this same study identified a three-factor model of the AQ with a Communication subscale distinct from Social Skills and Details and Patterns (as the diagnostic triad would predict), a pattern like that of the five-factor model arose when these three subscales were correlated against Big Five traits.

A replication of the same three-factor extraction revealed that whilst items loading on the first component were from the AQ Social Skills subscale, and items loading on the second component from the Details and Patterns subscale, the third component included item loadings from both the AQ Communication subscale and the AQ Social Skills and Details and Patterns subscales (Hurst et al., 2007). These patterns together support the idea that the *Communicative Deficits* symptom domain is an amalgam of social and non-social traits of distinct cognitive (and, perhaps, biological) origins. Finally, in each of the AQ-based factor analytic studies in which inter-factor correlation patterns were provided (Kloosterman et al., 2011; Stewart & Austin, 2009; Austin, 2009; Table 1), there were significant correlations between social interaction-related factors and communication factors, and between communication factors and inflexibility factors, but no such relationships between social interaction and inflexibility factors.

Evidence also supports the idea that the *Restricted and Repetitive Behaviors and Interests* factor of the diagnostic triad, like the *Communication Deficits* factor, conflates traits rooted in *Cognitive Inflexibility* (*e.g.,* restricted interests) and those rooted in lower-level *Sensory Abnormalities* (*e.g.,* auditory, tactile and visual hypersensitivities). Aforementioned ADI-R-based factor analyses focusing specifically on *Restricted and Repetitive Behaviors and Interests* diagnostic items all resulted in fairly clean splits between traits and behaviors concerning

cognitive inflexibility, and those concerning lower-level sensory issues (Szatmari et al., 2006; Papageorgiou et al., 2008; Lam, Bodfish, & Piven, 2008).

Two additional ADI-R factor analyses within the *Restricted and Repetitive Behaviors and Interests* element of the diagnostic triad likewise found separable factors of *Repetitive Sensory Motor Actions* and *Resistance to Change* in one case, and *Repetitive Sensorimotor* and *Insistence on Sameness* in the other (Cuccaro et al., 2004, and Richler, Bishop, Kleinke, & Lord, 2007, respectively). In a previously mentioned ADI-R analysis, three components - *Repetitive Motor Behaviors*, *Insistence on Sameness*, and *Circumscribed Interests* - were extracted from items subsumed under the *Restricted and Repetitive Behaviors and Interests* domain (Lam, Bodfish, & Piven, 2008). Even though *Circumscribed Interests* was identified as its own component, individual items pertaining to *Circumscribed Interests* were the only items loading on both *Repetitive Motor Behaviors* and cognitive flexibility-related *Insistence on Sameness* factors. Items loading on *Repetitive Motor Behaviors* and *Insistence on Sameness*, in other words, were distinct in their loadings. And the aforementioned ADI-R factor analysis by Tadevosyan-Leyfer et al. (2003) found distinct factors for inflexibility-based *Savant Skills* and *Sensory Aversions*, with a virtually nonexistent inter-factor correlation (Table 1). On the other hand, the *Compulsions* factor which, according to the current hypothesis, should have been be a heterogeneous mixture of inflexibility and sensory abnormalities, constituted an independent factor, with extremely low inter-factor correlations between these three factors. Finally, loadings from a factor analysis of the Childhood Autism Rating Scale (CARS) distinguished between items concerning fear, nervousness, and adaptation to change (i.e. reactions to novel situations reflecting cognitive flexibility) and those regarding sensory and motor issues related to taste, touch, smell, body use, and stereotypies (i.e. *Sensory Abnormalities*) (Magyer & Pandolfi, 2007).

The factor analyses of retrospective ADI-R and current ADOS-G inventories described above (Kamp-Becker et al., 2009) lend longitudinal support to the hypothesized heterogeneity of symptoms subsumed under the *Restricted and Repetitive Behaviors and Interests* domain (insofar as retrospective ASC trait inventories reliably reflect past trait presentation). Retrospective ADI-R items load onto 4 factors: *Social Communication*, *Anxiety and Compulsions*, *Stereotyped Behavior*, and *Inadequate Behaviors*; current presentation (ADOS-G) items loaded onto 5 partially coincident factors: *Social Communication*, *Non/Verbal Behavior*, *Hyperactivity*, *Stereotyped Behavior*, and *Circumscribed Interests/Compulsions*. Importantly, ADI-R items concerning circumscribed interests load on the *Anxiety and Compulsions* factor. A subset of the *Cognitive Inflexibility/Insistence on Sameness* factors in aforementioned studies distinguishes these interests from low-level sensory abnormalities; and the ADOS-G factor *Circumscribed Interests/Compulsions* maintains this separability from sensory abnormalities at current presentation. In the diagnostic triad, then, just as social reciprocity is distinct from cognitive flexibility within *Communicative Deficits*, sensory abnormalities are distinct from cognitive flexibility within *Restricted and Repetitive Behaviors and Interests*. These social/cognitive and sensory/cognitive distinctions pertain across retrospective (ADI-R) and current (ADOS-G) behavioral inventories.

Implicit in the distinction between *Sensory Abnormalities* and *Cognitive Inflexibilities* is the idea that abnormalities in sensory modulation are a primary, core feature of ASC, rather than secondary and largely peripheral to core ASC traits. As Rogers and Ozonoff (2005) point out, sensory issues have historically received only peripheral attention in explanatory accounts of autism, often subsumed under broader accounts such as executive functioning deficits and WCC. The reconceptualization of sensory modulation issues as primary to ASC is gaining traction in

the literature, however. Indeed, the most thorough meta-analysis to date found that sensory modulation issues consistently distinguish ASC and TD individuals across studies (Ben-Sasson et al., 2009). Sensory issues are largely separable not only from cognitive inflexibility (as argued above), but also from executive functioning (Boyd et al., 2009; Chen et al., 2009), and social interaction deficits (Watson et al., 2011), the latter observable as early as infancy (Baranek, 1999). Moreover, although the affected modalities can vary greatly between ASC individuals (perhaps explaining some of the ambivalence towards including sensory issues as a core symptom) there is some evidence that, within ASC individuals, extremity of sensory modulation abnormalities highly correlates between modalities (Kern et al., 2007).

Beyond the implications for specific diagnostic domain heterogeneities and the hypothesized alternative triad, the above analyses indicate more generally that social and non-social ASC traits are somewhat independent within individuals, contrary to the Kanner hypothesis. Social/non-social independence also arose in the form of factors *Social Affect* and *Restricted, Repetitive Behaviors* in a factor analysis of the ADOS, later replicated in an independent sample (Gotham et al., 2007; 2008). Such social/non-social separability was also the consensus among reviewed factor analyses of the AQ in more typically developing (i.e. sub- and non-clinical) individuals, with different analyses variously revealing: a) a diagnostic triad-like, three-factor model of *Social Deficits*, *Communication Deficits*, and (preference for) *Details and Patterns* (Austin, 2005, replicated in Hurst et al., 2007); b) a factor structure mirroring the five subscales originally intended by Baron-Cohen et al. during construction and validation of the AQ (*Social Skills, Communication, Restricted/Repetitive Behaviors, Imagination,* and *Attention to Detail* (Kloosterman, Keefer, Kelley, Summerfeldt, & Parker, 2011); and c) a four-factor model in-between these two, including *Imagination* but without *Restricted/Repetitive*

Behaviors (Stewart & Austin, 2009). Reconciling these discrepancies is a hierarchical factor structure with superordinate factors of *Social Interaction* and *Attention to Detail*, the former comprising four subfactors: *Social Skills, Attention Switching, Communication, and Imagination* (Hoekstra, Bartels, Cath, & Boomsma, 2008). The same hierarchical factor structure was replicated in another analysis (Valla et al., 2010); when analyzed separately by sex, the superordinate split between social and non-social traits was particularly strong in males, compared to females. In all these analyses, social factors were distinct from non-social factors.

Similar social/non-social separability has also arisen in other E-S correlations in typically developing individuals. For instance, resistance to visual contextual illusions (framed as an epiphenomenon of autism-like, detail-oriented processing biases) and Systemizing Quotient (SQ) scores have been shown to be independent of Empathizing Quotient (EQ) scores (Walter, Dassonville, & Bochler, 2008), while RMET scores are in one study of adults unrelated to EFT disembedding speed, performance on a block design test, and SQ score (Carroll & Chiew, 2006). Additional studies of empathizing-systemizing co-variance (discussed in further detail below) have found sex-dependent patterns wherein social and non-social ASC traits are more independent either in males compared to females (Voracek & Dressler, 2006), or females compared to males (Valla et al., 2010).

4.2.2 Evidence For the Kanner Hypothesis

Although the body of evidence countering the Kanner hypothesis is substantial, evidence supporting Kanner-like co-variance is not insignificant. Two analyses support a single factor model of ASC traits. Constantino and colleagues (2004) found principal components analyses of SRS and ADI-R inventories of PDD and non-PDD psychiatric cases supported a single underlying component accounting for a large proportion of individual variation in social and

non-social ASC traits (35% of variation in SRS scores, 40% in ADI-R scores). Although Szatmari et al. (2006) found a two-factor solution accounting for a majority of the variance (68% in a sample of 5y+ cases, 70% in a sample of 4-6y cases) in ADI-R scores, the first of these factors comprised severity of traits on all three subscales, with the second factor accounting for variance in overall level of functioning. Sample sizes for both studies were, however, notably smaller ($N = 226$, and $N = 129$, respectively) than most ADI-R factor analyses in the literature.

In some cases, Kanner-like co-variance has arisen alongside evidence against the Kanner hypothesis. In one of the ADI-R factor analyses aforementioned, two of three components extracted from *Restricted and Repetitive Behaviors and Interests - Repetitive Motor Behaviors* and *Insistence on Sameness* - correlated significantly and positively with social interaction deficits (Lam, Bodfish, & Piven, 2008). In the aforementioned endophenotypic analysis (Sacco et al., 2010), lower-level sensory abnormalities (e.g., decreased pain sensitivity) loaded primarily on the first of four factors (*Circadian and Sensory Dysfunction*) while more cognitively based repetitive behaviors and circumscribed interests (e.g., verbal stereotypies) loaded on the fourth factor (*Stereotypic Behavior*). This separability of high- and low-level repetitive behaviors reinforces the distinction between *Cognitive Inflexibility* and *Sensory Abnormalities* in the hypothesized alternative triad. There is also decreased variability within versus between monozygotic twin sibships in social and communicative deficits, and restricted interests, but not sensory issues (Kolevzon et al., 2004). Although this absence of detectable heritable co-variation does not support sensory symptoms' representing a primary factor in ASC, it does support the separability of *Sensory Abnormalities* from *Cognitive Inflexibility*.

In the twin studies that teased apart genetic and phenotypic overlap of the diagnostic triad, although each of the triad domains was highly, individually heritable, phenotypic overlap

between social and non-social domains was moderate but certainly greater than zero (Ronald et al., 2006a). (From the discussion of an alternative triad it could be argued that phenotypic overlap reflects the heterogeneity within diagnostic domains that the alternative triad is meant to correct; however, overlap existed even between *Social Interaction* and *Restricted and Repetitive Behaviors and Interests*, which have no symptoms in common even in the alternative triad.) Moreover, such phenotypic overlap increased as severity of social and non-social symptoms increased (Ronald et al., 2006b) – meaning that even if slight degrees of autistic traits in the typical population may be largely independent, more pronounced degrees of the same traits begin to reinforce each other across symptom domains, culminating in the apparent syndrome of autism.

Longitudinal studies support this notion of developmental reinforcement of symptoms across domains: although severity in diagnostic triad domains at 2y does not predict severity at 7y, severity of social and communicative deficits at 3y predicts all three diagnostic domains at 7y (Charman et al., 2005). Additional behavioral evidence for early emergence of social/non-social autistic trait co-variance includes replicated findings that RMET scores and visual perceptual disembedding speed on the EFT are inversely related in 4-5 year old children (Jarrold et al., 2000; Pellicano, Maybery, & Durkin, 2005). The former study found this co-variance pattern in children (9y) with ASC diagnoses as well, documenting some spectral continuity in such social/non-social co-variance. Later in development, poor ToM skills (measured via first and second order false belief tasks) are associated with both decreased susceptibility to visual illusions (in which detail-oriented processing gives one the advantage of ignoring visual contexts giving rise to such illusions) and increased scores on the Wechsler Block Design subtest, in a Kanner-like fashion, in typically developing adolescents and adults with subclinical autistic traits

(Best et al., 2008).

Similar Kanner-like social/non-social co-variance patterns in the normal adult population also arises between facial emotion labeling on an early version of the RMET and disembedding speed on the EFT (Jarrold et al., 2000). When re-analyzed separately by sex, however, this Kanner-like co-variance remains in women but not men. At the same time, specific patterns of sex-dependent co-variance vary from one study to another, with some finding co-variance in females but not males (Jarrold et al., 2000), others males but not females (Valla et al., 2010). In this latter study, men in systemizing-related undergraduate major fields were characterized more by empathizing deficits (measured via the RMET, the Benton Face Recognition Test, and the Social Interaction hierarchical subscale of the AQ) than by systemizing skills (measured via the EFT and the AQ Details/Patterns subscale), whereas no such link between systemizing field and face reading difficulties was present in females (Valla et al., 2010). Heterogeneity in sex-dependent co-variance findings is not restricted to the question of which sex exhibits the co-variance, either; it also seems to depend on the specific measures whose co-variation is being assayed. When co-variance is assayed from RMET scores and Systemizing Quotient scores, males but not females exhibit not inverse (Kanner-like) but actually *positive* empathizing-systemizing co-variance, with RMET scores predicting systemizing tendencies (Voracek & Dressler, 2006). Of these various sex-dependent outcomes, the pattern in Valla et al. (2010) is unique in that an independent study provides developmental support for it: Kanner-like co-variance has been found in school-age boys, but not girls (Skuse et al., 2009), supporting the "extreme male brain" notion that males are closer than females to the autism spectrum (Baron-Cohen, 2002).

4.3 Reconciling the Evidence: A Developmental Dynamic Interaction Model

The mutual exclusivity between social/non-social ASC trait independence and Kanner-like co-variance, an assumption implicit in many framings of the debate, downplays and overlooks the role of development in the progression of symptoms of autism – which is, after all, a developmental disorder. In so doing, the assumption overlooks a potential framework for reconciling what appear to be opposing sets of evidence. Independently heritable traits of ASC (i.e. the factors of the alternative triad) may give rise to Kanner-like co-variance in behavioral presentation as they dynamically interact and reinforce each other across development. It is in this developmental construction that non-social capacities can be drafted into the service of social cognitive ends, in the way that a positive correlation between figure disembedding ability and facial emotion reading has been interpreted as males applying systemizing skills to empathizing problems (Valla et al., 2010) (For experimental evidence of such piecemeal face processing, in ASC individuals, see Evers, Noens, Steyaert, and Wagemans, 2011). Explicit instruction of people with Asperger syndrome in solving empathizing problems using systemizing skills has been described as "systemizing empathy" (Golan & Baron-Cohen, 2006).

We argue that this phenomenon – the application of systemizing skills within cognitive domains that are more typically, more commonly or more efficiently the bailiwick of empathizing skills – need not be the subject of explicit instruction or explicit cognitive strategy, and is not restricted to people with autism-spectrum conditions. We argue, rather, that it happens during typical cognitive development as systemizing and empathizing skills supplement – and in some instances supplant – each other during interaction with environmental task demands.

Indeed, reconsidering the reviewed evidence through this developmental lens helps clear up some seeming contradictions amongst findings within and between studies. A prime example

of how the hypothesized developmental model would manifest empirically can be found in a series of related studies on a single large twin sample (Ronald et al., 2006a, b; Dworzynski et al., 2007). First, aforementioned findings demonstrate the non-mutual-exclusivity of the two sides of the Kanner hypothesis debate, supporting Kanner-unlike, independent heritabilities of social and non-social ASC traits, but also Kanner-like, moderate correlations between domains of behavioral presentation (Ronald et al., 2006ab). Correlations between diagnostic domains of *Communicative Deficits* and both *Social Interaction Deficits* and *Restricted and Repetitive Behaviors and Interests* can be ascribed partly to the heterogeneity in the communication domain, argued above. Such intra-domain heterogeneity within the *Communicative Deficits* domain would give the appearance of deeper inter-domain links than the hypothesized possibility: that social interaction deficits are incorrectly paired with cognitive flexibility-related linguistic deficits affecting communication. Heterogeneity within domains cannot account for correlations between the distinct diagnostic domains *Social Interaction Deficits* and *Restricted and Repetitive Behaviors*, however.

The same data from this large twin sample were also reanalyzed separately within subgroups of individuals who presented with extremity (top 5% cutoff) in one, two, and three symptom domains. Behavioral overlaps between symptom domains (i.e. intra-domain correlations of variation in symptom severity) were significantly greater within the group of individuals exhibiting severity in all three domains, compared to groups of individuals exhibiting top 5% severity in zero or one domain (Ronald et al., 2006b). Similar increased overlaps related to extremity in multiple domains have been found in other studies (*e.g.,* Skuse et al., 2009; Frazier et al., 2008). According to Baron-Cohen's "extreme male brain" notion this heightened correlation at the extreme should be doubly true of males, who are closer to the autism spectrum;

and indeed, the same study found greater inter-domain behavioral overlap in males than females. Greater Kanner-like co-variance in males compared to females is a common albeit not a universal result in studies of autistic traits (Valla et al., 2010). The hypothesized developmental model may also support the conjecture that a positive correlation in males between figure disembedding and facial emotion labeling (EFT and RMET, respectively) reflects a predominantly male cognitive strategy of applying systemizing skills to empathizing problems (Valla et al., 2010).

The present model of developmental interaction between symptom domains is consistent with longitudinal observations of the aforementioned twin samples (Ronald et al. 2006ab), measured by the MacArthur Communicative Development Inventories at 2, 3, and 4 years, and the CAST at 8 years. Early language deficits and later *Social Interaction* and *Communication Deficits* were related primarily genetically, rather than phenotypically. Although early language deficits were genetically unrelated to later *Restricted and Repetitive Behaviors and Interests*, *Communication Deficits* and *Restricted and Repetitive Behaviors and Interests* domains were phenotypically related later in development. A combination of independent heritability and phenotypic dependence arising in concert as they did in these twin samples supports developmentally arising Kanner-like co-variance (even after taking into consideration the shared cognitive inflexibility component), with later *Communicative Deficits* representing a developmentally emergent interaction between cognitive inflexibility and linguistic elements of social behavior.

This model of *Communicative Deficits* emerging developmentally as primary social reciprocity- and cognitive inflexibility-related factors interact is also reflected in the aforementioned BISCUIT-1 factor analysis (Matson et al., 2010). In the earlier developmental

stage represented by this study's sample, the *Communicative Deficit* factor seems to have begun emerging, carrying a small number of inventory items, while many reciprocation-related communicative behaviors load on a *Socialization/Non-verbal Communication* factor, and verbal stereotypies load with cognitive inflexibilities and sensory issues on a *Repetitive Behaviors/Restricted Interests* factor. This study also highlights the idea, distinguishing the current model from previous models, that a secondary deficit, though it emerges from more primary behavioral interactions, can take on a 'life of its own', its presentation becoming independent of the primary factors from which it emerges.

The proposed developmental model helps explain other seeming empirical contradictions. The longitudinal finding that communicative deficits are not predictive of any symptom domain at 2y, but significantly predict all three diagnostic domains later in childhood (Charman et al., 2005) would be expected if inter-domain behavioral interactions increase with development. At 4-5y verbal rituals load on the same factor as repetitive use of objects and other manifestations of cognitive inflexibility (Kamp-Becker et al., 2009). By 6-7y, a communication-related factor emerges encompassing stereotyped phrases, eye contact, and facial expressions. The hypothesized model would explain that whereas social deficits may be largely independent of cognitive inflexibility early in development, behavioral and cognitive interactions over time produce repetitive phrases - part linguistic inflexibility, part social communicative deficit.

In another apparent contradiction, the loading of "age of first social smile" on a *Stereotypic Behavior* factor (Sacco et al., 2010) may seem to suggest early links between *Social Interaction Deficits* and *Restricted and Repetitive Behaviors and Interests* diagnostic domains, as in single-cause accounts of autism (*e.g.,* Constantino et al., 2004). In the context of the hypothesized alternative triad and developmental model, however, one might posit that "age of

first social smile" is an early example of behavioral interactions between *Social-Communicative Deficits* and stereotypy-related *Cognitive Inflexibility* factors. In case cognitive flexibility may seem behaviorally unrelated to social smiling, consider the rapid cognitive control in play during a smiling response to a social overture, or during a decision to share an emotional response with a social partner at the same time as one is busy experiencing that response.

Table 8. Summary of ASC Trait Factor Analytic Studies Reviewed

Study	Traits	Sample	Age *M* (*SD*) Range	Measure	Factor Rotation Method	Factors	% Variance Explained (EFA)
Matson et al. (2010)	All	405 ASC, PDD	2.26 (0.4) 1.26-3.08	BISCUIT-Part 1	Oblique (Promax)	Socialization/Non-verbal Communication	33.2%
						Repetitive Behaviour/Restricted Interests	
						Communication	
Magyer & Pandolfi, 2009	All	164 ASC, PDD	3.61 (1.65) 1.6-6.8	CARS	Orthogonal (Varimax, Quartimax), Oblique (Promax, Direct Oblimin)	Social Interaction	57.16% (Orthogonal) 41.67% (Oblique)
						Social Communication	
						Emotional Regulation	
						Stereotypies and Sensory Abnormalities	
Georgiades et al., 2007	All	209 ASC, PDD	5.46 (2.25) 2.3-40.2	ADI-R	Orthogonal (Varimax)	Social Communication	50%
						Inflexible Language and Behavior	
						Repetitive Sensory and Motor Behavior	
Szatmari et al., 2002	All	129 ASC, PDD	5.5(0.9) 4-6y 11.6(5.8) 5-N/A	ADI-R	Orthogonal (Varimax)	Symptom Severity	70.3% 67.7%
						Overall Level of Functioning	
Gotham et al., 2007	All	1139 ASC, PDD	5.64 (2.2) 1.2-15.3	ADOS	Oblique (Promax)	Social Affect	N/A
						Restricted and Repetitive Behaviors	
Gotham et al., 2008	All	1259 ASC, PDD	6.2 (2.12) N/A	ADOS	Oblique (Promax)	Social Affect	N/A
						Restricted and Repetitive Behaviors	
Snow et al., 2009	All	1861 ASC, PDD	8.32 (3.16) 4-18	ADI-R	Oblique (Quartimin)	Social Communication	N/A
						Restricted/Repetitive Behaviors	
Frazier et al., 2008	All	1170 ASC, PDD	9 (4.88) 2-46	ADI-R	Oblique (Promax)	Social Communication	N/A
						Restricted/Repetitive Behaviors	
Boomsma et al., 2008	All	263 ASC, PDD	11 (5) 4-24	ADI-R	N/A	Play Skills	N/A
						Social Communication	
						Stereotypies in Communication and Behavior	
Van Lang et al., 2006	All	255: 125 TD, 130 ASC, PDD	11.25 (3.92) 4-20	ADI-R	N/A	Play Skills	N/A
						Social Communication	
						Stereotypies in Communication and Behavior	

RRBI = Restricted and Repetitive Behaviors and Interests; TD = Typically Developing; ASC = Autism Spectrum Conditions; PDD = Pervasive Developmental Disorders; CARS = Childhood Autism Rating Scale; ADI-R = Autism Diagnostic Interview – Revised; ADOS = Autism Diagnostic Observation Schedule; AQ = Autism Spectrum Quotient

Table 8., Continued

Study	Traits	Sample	Age *M* (*SD*) Range	Measure	Factor Rotation Method	Factors	% Variance Explained (EFA)
Kamp-Becker et al., 2009	All	140: 35 TD, 105 ASC, PDD	12.04 (4.59) 6-24	ADI-R	N/A	Social Communication	N/A
						Inadequate Behaviors	
						Anxiety and Compulsions	
						Stereotyped Behavior	
Tadevosyan et al., 2003	All	292 ASD	15.58 (6.21) 2-47	ADI-R	Oblique (Varclus)	Spoken Language	41%
						Social Intent	
						Compulsions	
						Developmental Milestones	
						Savant Skills	
						Sensory Aversions	
Constantino et al., 2004	All	226 ASD, PDD, TD	4-18y	ADI-R	Orthogonal (Varimax)	Single factor	35%
Hurst et al., 2007	All	1005 TD	19.36 (3.89) 17-55	AQ	Oblique (Promax)	Social Skills	29%
						Communication	
						Attention to Detail	
Valla et al., 2010	All	144 TD	20.2 (1.86) 18-27	AQ	Orthogonal (Varimax)	Social Interaction	64%
						Preference for Details/Patterns	
Kloosterman et al., 2011	All	522 TD	21 (5.15)	AQ	Oblique (Promax)	Social Skills	45%
						Communication	
						Imagination	
						Attention to Detail	
						Restricted and Repetitive Behaviors	
Stewart & Austin, 2009	All	536 TD	24.3 (10.5)	AQ	Orthogonal (Varimax)	Social Skills	29%
						Communication	
						Imagination	
						Attention to Detail	
Hoekstra et al., 2008	All	1374: 1338 TD, 15 ASC, 15 OCD, 15 SAD	24.65 (4.32)	AQ	Oblique (Promax)	Social Interaction	N/A
						Preference for Details/Patterns	
Austin, 2005	All	337 TD	34.22 (7.23)	AQ	Oblique (Promax)	Social Skills	28%
						Communication	
						Attention to Detail	

RRBI = Restricted and Repetitive Behaviors and Interests; TD = Typically Developing; ASC = Autism Spectrum Conditions; PDD = Pervasive Developmental Disorders; CARS = Childhood Autism Rating Scale; ADI-R = Autism Diagnostic Interview – Revised; ADOS = Autism Diagnostic Observation Schedule; AQ = Autism Spectrum Quotient

Table 8., Continued

Study	Traits	Sample	Age *M* (*SD*) Range	Measure	Factor Rotation Method	Factors	% Variance Explained (EFA)
Richler et al., 2007	RRBI	279: 65 TD, 165 ASC, PDD	2.23 (0.45)	ADI-R	Orthogonal (Varimax)	Insistence on Sameness Repetitive Sensory and Motor Behavior	N/A
Papageorgiou et al., 2008	RRBI	153 ASC, PDD	5.96 (3.3) 1.6-19	ADI-R	Orthogonal (Varimax)	Insistence on Sameness Repetitive Sensorimotor Behaviors, Interests	52%
Szatmari et al., 2006	RRBI	339 ASC, PDD	8.4 (5.51)	ADI-R	Orthogonal (Varimax)	Insistence on Sameness Repetitive Sensory and Motor Behavior	36% (current), 33% (ever)
Lam et al., 2008	RRBI	361 ASC, PDD	9.02 (6.15) 1.6-29	ADI-R	Orthogonal (Varimax)	Insistence on Sameness Circumscribed Interests Repetitive Motor Behavior	52%
Cuccaro et al., 2003	RRBI	207 ASC, PDD	9.06 (4.57) 2.4-21.2	ADI-R	Oblique (Promax)	Resistance to Change Repetitive Sensory and Motor Behavior	32%

RRBI = Restricted and Repetitive Behaviors and Interests; TD = Typically Developing; ASC = Autism Spectrum Conditions; PDD = Pervasive Developmental Disorders; CARS = Childhood Autism Rating Scale; ADI-R = Autism Diagnostic Interview – Revised; ADOS = Autism Diagnostic Observation Schedule; AQ = Autism Spectrum Quotient

4.3.1 Applying the Proposed Developmental Model: A Novel Account of Repetitious Behaviors

In the proposed developmental model, primary, initially independent *Social Interaction Deficits*, *Cognitive Inflexibility*, and *Sensory Abnormalities* behaviorally interact and become related. Here we demonstrate the utility of this model by explaining the non-primary autistic trait of repetitive behaviors in terms of interactions between the latter two primary traits. To be clear, this is not meant to imply that these two primary traits interact more or less than any other pairing; to highlight this point, we follow our discussion of repetitive behaviors with a brief account of how communicative deficits would arise from interactions between *Social Interaction Deficits* and *Cognitive Inflexibility*.

In the current diagnostic triad, repetitive behaviors are considered primary symptoms, and

are subsumed under the domain of *Restricted and Repetitive Behaviors and Interests*; sensory abnormalities, meanwhile, are treated as secondary symptoms. On face, repetitive behaviors are more easily categorized and accounted for using the diagnostic triad than the alternative triad, as such behaviors do not fit as clearly under any of the alternative triad domains. Viewing the alternative triad through the lens of the proposed developmental model, however, such behavioral stereotypies can be explained as external, motoric manifestations of developmentally snowballing interactions between independent primary traits of *Cognitive Inflexibility*, and *Sensory Abnormalities* such as hypersensitivity to auditory, visual, or tactile stimuli. This would explain why Chen et al. (2009) found that sensory hypersensitivities and cognitive inflexibility each predict the presence of repetitive behaviors, yet are unrelated to each other. Repetitious movements may be externalized efforts to impart predictability on a world rendered chaotic and intractable by an abnormal, cognitively inflexible, perceptually overstimulated mental environment (Belmonte, 2008). Relationships between ritualistic behaviors and perceived uncertainty or fearfulness exist likewise in young, typically developing children (Evans et al., 1999).

A crucial distinction, demonstrating the advantages of the present developmental interactionist account, is that as repetitive behaviors are continually used to compensate for cognitive-sensory interaction chaos, the more likely it becomes that neural networks hard-coding such interactions will arise. Once encoded in dedicated neural networks, these behaviors could then take on a 'life of their own', arising under conditions that may not have triggered them prior to developmental snowballing. Repetitive movements might arise in part as externalized efforts to impart predictability to a world rendered chaotic and intractable by an abnormal, cognitively inflexible and perceptually overstimulated mental environment (Belmonte, 2008). Overall level

of functioning, then, may determine the executive control capacity available to inhibit externalizing, overtly motor behaviors. Every individual would have some threshold beyond which cognitive-perceptual load could not be pushed without evoking externalizing behaviors. High-functioning people, with high thresholds, when confronted with sensory perceptual challenges would manifest mostly internalized responses that impose conceptual 'sameness,' whereas in low-functioning individuals this drive to render perceptual organization tractable might more often spill over into externalized, motor stereotypies that directly, mechanically impose sensory sameness.

Multiple findings support this characterization of repetitive behaviors as emergent from an interaction of more primary sensory and cognitive symptoms, with overall level functioning determining the internal or external manifestation of such interactions. Positive relationships between IQ and age, and restricted interests, have arisen in some cases, as well as negative relationships between IQ and age, and sensorimotor stereotypies (Szatmari et al., 2006; Papageorgiou et al., 2008). (IQ controlled for age, and age, within individuals, are taken here as correlates of overall functioning). Likewise, factor analyses of retrospective (childhood) ADI-R items in one of the aforementioned longitudinal studies (Kamp-Becker et al., 2009) reveal an *Anxiety and Compulsions* factor; if internalized anxieties arise from cognitive inflexibilities faced with overstimulating novel situations, compulsive behaviors would be their external analogues. The event-related timing of dorsolateral prefrontal cortical activation in a non-social visual attention task in boys with autism spectrum conditions correlates with ADI-R social and communicative subscores but not with the ADI-R repetitive behaviors subscore (Belmonte et al., 2010), as might be expected if repetitive behaviors were an ætiologically secondary symptom.

On the ADOS-G measure of current symptom presentation in the same participants, items

related to anxiety loaded negatively on the ADOS-G factor *Hyperactivity*, meaning that decreased hyperactivity was related to increased anxiety. Such a relationship would be expected if internalized manifestations of cognitive inflexibility and hypersensitivity to context novelty (e.g., anxiety in novel situations) increase as externalized manifestations (e.g., hyperactivity) decrease (Kamp-Becker et al., 2009). The same investigation also found that full-scale IQ was at once unrelated to the *Social Communication* factor on both retrospective (childhood) ADI-R and current ADOS-G inventories, and significantly, negatively correlated with behavioral stereotypies at both time points.

In this respect, some differences between ADI-R analyses may be complementary rather than contradictory. Where one analysis (Szatmari et al., 2008) found evidence for influence of higher cognitive functioning (as indicated by IQ and age) on ASC symptoms, in the form of a cognitively based need for sameness, others found evidence for the symptomatic manifestation of lower cognitive functioning in ASC, in the form of ritualistic behaviors (Lam et al., 2008; Kamp-Becker et al., 2009); one of the latter found this relationship both in early development and later presentation (Kamp-Becker et al., 2009). When extracted factors from the *Restricted and Repetitive Behaviors and Interests* diagnostic domain were correlated with age, Vineland Adaptive Behavior Scale scores (VERBS), and ADI-R domain scores, *Insistence on Sameness* presentation (current and ever) was significantly, positively correlated with the ADI-R communication domain, and *Repetitive Sensory and Motor Behaviors and Interests* also were highly negatively correlated with adaptive functioning and IQ, measured via the VERBS and Loiter IQ, respectively. The former result provides additional support for the proposed alternative triad in which communicative inflexibilities are grouped with the high-level repetitive behavior of circumscribed interests, as in Georgiades et al. (2007). The latter provides more

concrete support for the idea that overall level of functioning (in this case indicated by IQ) may determine, or may correlate with, the threshold between internalized, cognitive drives, and low-level repetitive behavior of externalized, stereotypic, motor drives.

Of course, these are inferences about divergent developmental trajectories based on cross-sectional data; even the longitudinal analysis (Kamp-Becker et al., 2009) relied on retrospective ADI-R inventories for inferences about early developmental patterns. There is some more direct developmental support for this hypothesis, though; a within-subjects, repeated measures, longitudinal analysis comparing high- and low IQ (median split) groups from age 4 through age 19, with an intermediate data collection at age 13, using the ADI-R (with an added emotional responsiveness portion), ADOS, and the Vineland Adaptive Behavior scale (VABS), as well as the Stanford-Binet IQ test (McGovern & Sigman, 2005), found that both IQ groups significantly improved in *Social Interaction Deficits* and *Restricted and Repetitive Behaviors and Interests* domains with age, but that the high IQ group improved significantly more than the low IQ group in *Restricted and Repetitive Behaviors and Interests*, and not in *Social Interaction Deficits*. Thus, across development higher functioning carries advantages specific to reducing repetitive behaviors.

The novel account of repetitive behaviors offered here reconciles the multiple hypotheses of the origins and functioning of such behaviors offered by Turner (1999). In imparting predictability on cognitive/perceptual chaos arising at the junction of *weak central coherence*-related cognitive inflexibility, and sensory hypersensitivity, such behaviors serve as a *homeostatic mechanism* at the sensory perceptual level. If the over-arousing stimulus is another person whose movements/touch/etc. cannot be predicted, these behaviors avoid reliance on the *impaired metalizing* that would otherwise render the other person's actions predictable. Once

such behaviors take on a life of their own in interactively specialized neural networks, they become *operant behavior*, present even in the absence of what originally brought them on. Such neural networks would also make it more difficult to inhibit these behaviors, or to extinguish them via generation of novel behavior, both of which symptoms appear as aspects of *executive functioning* deficits.

Meanwhile, inconsistencies in relating level of functioning to internalized and externalized manifestations of 'sameness' are reconciled by the notion that overall level of functioning establishes a maximum cognitive capacity for integrative processing and abstraction, determining whether sensory inputs are internalized or externalized. As Turner notes, high-functioning people with ASC can exhibit 'low-level' motoric repetition, and low-functioning people with ASC can exhibit an internalized insistence on sameness. Such apparent inconsistencies can be accounted for by the present threshold model of overall functioning: sensory inputs hyper-arousing enough to surpass high functioning individuals' greater threshold could produce externalized motoric compensation, whilst some inputs beneath even the lower threshold of low functioning individuals could produce an internalized compensatory need for sameness.

The hypothesized relationship between cognitive inflexibilities and compensatory repetitive behaviors is well summarized in an interpretation of sibling correlations of ASC symptom severity (Spiker, Lotspeich, Dimiceli, Myers, & Risch, 2002): "while all the individuals with autism had some ritualistic or repetitive behaviors or preoccupations, with increasing cognitive ability these tend to be more symbolic and less motoric." Just as typical cognitive development elaborates iconic sensory and immediate motor capacities into symbolic concepts and sequential, goal-directed plans, atypical cognitive development elaborates a need

for sameness in these concrete domains into a need for scripting in domains that typically are more abstract. Social psychology in typical individuals has demonstrated that the low "level of construal" of relatively immediate, unprocessed, unelaborated, detail-oriented, iconic perception evokes a similarly low "psychological distance" from one's own immediate perspective on objects, events, contexts, and social partners (Trope & Liberman, 2010), giving rise to a cognitive style that emphasizes peripersonal rather than distant spaces, current rather than past or future events, absolute rather than conditional plans, and egocentric rather than allocentric social perspectives – and *vice versa,* low psychological distance evokes low level of construal.

Such effects arise during short-term, situational manipulations; analogous long-term, developmental influences might be expected to be even more potent. As such, it is not difficult to conceive of how *Cognitive Inflexibilities* in higher-functioning ASC individuals might interact with, and developmentally adapt to, a basic *Sensory Abnormality* such as auditory hypersensitivity in such a way that narrow specialties such as musical savantism may emerge in the factor space between these two more primary, independently inherited traits. Or, for a less exceptional example, the *Communicative Deficits* domain of the diagnostic triad may be better understood as deficits emerging developmentally from the factor space between primary *Social Interaction Deficits* and *Cognitive Inflexibility* domains of the alternative triad: whilst the former domain impedes development of the back-and-forth conversational reciprocity, the latter domain impedes the linguistic and conceptual fluidity and flexibility needed to adapt as conversations unfold and progress in unpredictable, unscripted ways.

4.4 The Bigger Picture

The advantage of a developmental perspective may seem obvious; behavioral theorists are well inured to the idea that developmental models are essential for teasing apart nature and

nurture. A previous review addressing the Kanner hypothesis, for instance, argues that the diagnostic triad is fractionable in terms of causal origins, with a concluding caveat that ASC traits do in any case interact developmentally (Happé & Ronald, 2008). The developmental model posed here is more than an elaboration of such developmental caveats, though, offering an explanation for Kanner co-variance, but also how secondary, compensatory behaviors like repetitive behaviors can become stable aspects of autistic behavior, their expression dissociated from their original coping purpose.

Specifically, the dynamic, developmentally snowballing interactionism posited here is best described within the model of "interactive specialization" (Johnson, Halit, Grice, & Karmiloff-Smith, 2002). A typical developmental perspective would take the view that independently heritable social-communicative deficits and cognitive inflexibilities both are programmed into neural architecture and begin to co-vary as they are expressed through shared behavioral and cognitive outlets. An interactive specialization view, on the other hand, would hold that as these traits are integrated in co-expression repeatedly across development, their integration would come to be reflected in neural networks. These altered networks, in turn, become co-conspirators in this process, encouraging further integration by serving as a path of increasingly less neural resistance.

If development occurred according to the former model, it would not be as clear why social and cognitive integration would become increasingly deeper to the extent that observations like Kanner's depict restricted interests as a concomitant of social dysfunction. Within the interactive specialization view, however, it seems more likely that a bio-behavioral mechanism that produces social aversion may also induce an equal and opposite reaction of object preoccupation. A child with ASC may spend early years being repeatedly bombarded with the

unsolicited advances of what appear to be large objects with irreproducible, inexplicable cause-and-effect. Meanwhile, the more physically salient details of the inanimate world capture their attention: a spinning ceiling fan is a more visually salient motion than a moving mouth, a car horn a more intense sound than a voice. Physiologically, autism may begin with pathologically low-entropy local neural networks; these aberrant local representations then prevent the normal activity-dependent development of long-range, integrative connections (Belmonte et al., 2004ab) via a bio-behavioral feedback loop. As the child develops, they repeatedly retreat from the unpredictability of people to the predictability of the inanimate world, and their attention is captured by small parts of this inanimate world. As with all behavioral skills and strategies, practice makes perfect: repetition of these learnt coping behaviors can allow underused networks to atrophy, specializing the brain for a detail-over-context, iconic-over-symbolic, egocentric-over-allocentric mode of perception and cognition.

In this view, the hypothesized developmental model, and interactive specialization more generally, seem more equipped to explain autism as a *developmental* disorder, and to account for the Kanner-like co-variance observed to arise between initially independent traits over the course of development. This paradigm may aid behavioral scientists appealing for genetic links to behavioral outcomes, as well as cognitive neuroscientists interested in tying the proposed model to neural mechanisms

4.4.1 Future Directions

Based on the present review, the most important theoretical shift to be made is to recognize that debates over single versus multiple causes of autism mean different things at the genetic, neurological and behavioral levels, and that a dynamic developmental lens is essential to understanding autism across these three levels of analysis. If, for instance, a single genetic effect

leads to low entropy local networks, a single 'cause' could look multi-factorial from the outset at the neurological level, each disturbed inter-regional connection an additional neurological 'cause', with the behavioral result being initially independent effects on social and non-social behavior. These behavioral effects would then interact developmentally according to the proposed model, leading to networks interactively specialized for the collection of behavioral traits we call autism, with social and non-social behaviors, and their neurological underpinnings, being somewhat connected, somewhat independent, and highly individualized. In a different scenario, if social and non-social ASC traits were to begin entirely independent at both genetic and neural levels, interactive specialization could result in their being seemingly related behaviorally and neurologically downstream.

The wide age ranges in the reviewed studies – varying as widely as 2-46y, frequently with maximum ages greater than 20y, and mean ages greater than 8y – present a problem of interpretation. For the present synthesis, inferences about developmental trajectories of social/non-social co-variance could be made only from three sources: the few longitudinal investigations that currently exist (e.g., Kamp-Becker et al., 2009); cross-sectional investigations repeating their analyses separately for younger and older sub-sections of the total sample (e.g., Szatmari et al., 2002); and meta-analytic comparisons of studies at early ages (e.g., Matson et al., 2010) to other, less restricted samples. It is only through the groundwork laid by such studies that the full importance of a developmental picture can be seen; with wide age ranges come large sample sizes (e.g., Frazier et al., 2008). We also sympathies with the fact that every subject counts when it comes to such analyses. As such, rather than excluding older cases, we encourage investigators factor-analyzing ASC traits to include developmental elements. Entering age as a co-variate in regression models, or segmenting samples cross-sectionally into age groups for

separate factor analyses are two straightforward ways to add a developmental element. In the latter case, cross-sectional cuts between early ages would ideally be as thin as statistical power would allow, to capture as much early behavioral dynamism as possible.

A thorough test of the model proposed here would involve longitudinally tracking the developmental trajectory of relationships between ASC trait domains from infancy through late childhood. If the aforementioned account of repetitive behaviors is correct, then the fact that such behaviors frequently arise early in development implies that inter-trait interactions rapidly produce secondary traits. Particular emphasis would be placed on potential loci of inter-domain interaction (e.g., joint attention), and aspects of behavior hypothesized to be emergent manifestations of inter-trait interactions. In tracking repetitive behaviors, for instance, it would be important to note the contexts in which they initially arise (e.g., in response to loud noises), as well as when they began appearing in the absence of the initially associated context; the former would indicate the compensatory, primary trait associations of repetitive behaviors, and the latter when, developmentally, repetitive behaviors become an independently developing trait domain.

The ideal instrument for such a study would be a developmentally-minded behavioral inventory, as extensive as the ADI-R or ADOS-G but, unlike these inventories, allowing for re-testing at different ages, to document changes in trait severity, and developmental emergence of secondary traits, including the contexts in which, for instance, repetitive behaviors arise (e.g., in response to loud noises). Constructing and validating inventories has been a major accomplishment for the field; what is needed now is the capability to apply these inventories to tracking age-related changes in autism. To date, the closest that studies have come to this approach are current-versus-ever comparisons of the ADI-R (Kamp-Becker et al., 2009); and a combination of multiple data collection points (2, 3, and 4 years old) using the MacArthur

Communicative Development inventory, and the CAST at age 8 (Ronald et al., 2006ab). However, the 'ever' component of the former is susceptible to hindsight biases and recall accuracy; and although the latter tracks early developmental changes in communication using multiple data points on the MacArthur, this inventory only measures communication traits, and transitioning to the CAST to inventory later childhood behaviors raises issues of inter-instrument consistency.

If the ideal instrument were available, exploratory factor analyses, treating inventory items individually, could then be compared between data collection points. The proposed model would predict that between infancy and early childhood, factor analyses would indicate primary *Sensory Abnormalities, Cognitive Inflexibility,* and *Social Interaction Deficits* factors. These would interact, and by late childhood communicative deficits would arise and load on the latter two factors; and secondary repetitive behaviors would emerge and load on the former two factors. By adolescence, Kanner-like behavioral interactions between factors will become more neurologically stable, resulting in emergent *Communicative Deficits* and *Repetitive Behaviors* factors, reflecting behavioral presentation that is partly independent of primary factor associates. If such developmental comparisons were repeated for individuals from the bottom 85%, 85-95%, and top 5% of inventory scores, then the degree of Kanner co-variance should be more pronounced with increased trait extremity. *Communicative Deficits* and *Repetitive Behaviors* factors would arise earliest in the top 5% group, latest (if at all) in the bottom 85% group.

The most productive shift for diagnosis and treatment would be to institute a new core diagnostic triad of primary behavioral symptoms – *Social Interaction Deficits*, *Cognitive Inflexibility*, and *Sensory Abnormalities* – that better reflects the factor structures found in the literature. The proposed DSM-V revisions to the diagnostic criteria for autism are a start. For

instance, the proposed changes include moving reciprocation-related behaviors from the current *Communicative Deficits* factor into a broader social-communicative deficit category, and moving verbal inflexibilities into a broader restricted and repetitive behaviors and interests category. Further progress, then, can be made by recognizing the heterogeneity of behaviors categorized as restricted and repetitive behaviors and interests, separating cognitive inflexibilities and sensory dysfunction into distinct factors. The diagnostic definition of autism should also present symptoms in a hierarchical fashion, distinguishing between a primary triad of *Social Interaction Deficits*, *Cognitive Inflexibility*, and *Sensory Abnormalities* and secondary symptoms (e.g. repetitive behaviors). In this way, the presence of common secondary symptoms could be treated as an indication of co-morbid primary traits, but diagnosis would not depend on symptoms that manifest in highly individualized ways.

If Kanner co-variance does emerge through interactions between distinct social and non-social traits, then non-social behaviors arising from *Cognitive Inflexibilities* might be exploited for social cognitive therapeutic ends – a "back door" route to training social cognition (Belmonte et al., 2010). Boyd, Conroy, Mancil, Nakao, and Alter (2007) is a perfect demonstration of the behavioral mechanism through which Kanner-like co-variance arises, as well as the potential therapeutic benefits of exploiting this mechanism: Peer-initiated play interactions are briefer when the toys in question lie outside the circumscribed interests of ASC children than when play is initiated with toys relevant to circumscribed interests. When toys of interest are present, ASC children's initial social bids to peers occur faster than when less preferred toys are present. With the right context, attention to social interaction might be bootstrapped via attention to special interests.

4.5 Conclusion

Biologically and behaviorally emergent aspects of autistic traits and their sub- and non-clinical manifestations can both be better understood as an independently or semi-independently inherited triad of primary factors *Social Interaction Deficits*, *Cognitive Inflexibility*, and *Sensory Abnormalities* dynamically interacting over time in cognition and behavior, giving rise to the quintessentially autistic inverse relationship between social and non-social skills first described by Kanner. This dynamic model of autistic development provides theoretical insights into the nature and nurture of autism and, in so doing, informs more effective treatment, too. Emergent symptoms, though diagnostically valid, would not be ætiologically primary (Yoder & Belmonte, 2011), and early therapies addressing primary symptoms (*Cognitive Inflexibility*, *Sensory Abnormalities*) might be a more effective treatment for emergent symptoms than are therapies that attempt to confront behavioral diagnostic symptoms head-on. For instance, early non-social cognitive exercises targeted at executive control and integration could exert knock-on effects on later social cognition (Belmonte et al., 2004ab). This conjectured developmental mechanism would reinforce the case for the earliest possible intervention, and further would target such intervention at aetiologically primary rather than diagnostically determinative symptom domains. The aim of these early therapies would be to work with, rather than against, autistic brain physiology and autistic cognitive style (Chen et al., 2012), to channel development towards more integrative sensory and cognitive control.

REFERENCES

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the autism spectrum quotient (AQ). *Personality and Individual Differences, 38,* 451–460.

Baranek, G. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months of age. *Journal of Autism and Developmental Disorders, 29, 213–*224.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences,* 6, 248–254.

Baron-Cohen, S., & Hammer, J. (1997a). Is autism an extreme form of the male brain? *Advances in Infancy Research, 11,* 193–217.

Baron-Cohen, S., & Hammer, J. (1997b). Parents of children with Asperger syndrome: What is the cognitive phenotype? *Journal of Cognitive Neuroscience, 9,* 548-554.

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,* 163–175.

Baron-Cohen, S., Hoekstra, R., Knickmeyer, R., & Wheelwright, S. (2006). The autism-spectrum quotient (AQ) – adolescent version. *Journal of Autism and Developmental Disorders, 36,* 343-350.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of

mind? *Cognition, 21,* 37–46.

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

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,* 361-374.

Baron-Cohen, S., Wheelwright, S. (2004). The Empathy Quotient (EQ). An investigation of adults with Asperger Syndrome or High Functioning Autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163-175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. G., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males, females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.

Baron-Cohen, S., Wheelwright, S., Hill, J, Raste, Y, & Plumb, I. (2001). The 'Reading the mind in the eyes' test revised version: A study with normal adults, and adults with Asperger Syndrome or High-Functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241-252.

Belmonte, M.K. (2008). Human, but more so: what the autistic brain tells us about the process of narrative. In: *Autism and Representation* (Mark Osteen, ed.), pp 166-179. New York: Routledge.

Belmonte, M.K., Allen, G., Beckel-Mitchener, A., Boulanger, L.M., Carper, R.A., Webb, S.J. (2004a). Autism and abnormal development of brain connectivity. *Journal of*

Neuroscience, 24, 9228-9231.

Belmonte, M.K., Cook, E.H. Jr., Anderson, G.M., Rubenstein, J.L., Greenough, W.T., Beckel-Mitchener, A., Courchesne, E., Boulanger, L.M., Powell, S.B., Levitt, P.R., Perry, E.K., Jiang, Y.H., DeLorey, T.M., & Tierney, E. (2004b). Autism as a disorder of neural information processing: directions for research and targets for therapy. *Molecular Psychiatry, 9,* 646-663.

Belmonte MK, Gomot M, Baron-Cohen S. (2010). Visual attention in autism families: 'unaffected' sibs share atypical frontal activation. *Journal of Child Psychology and Psychiatry, 51,* 259-276.

Ben-Sasson, A., Hen, L., Fluss, R., Cermak, S. A., Engel-Yeger, B., & Gal, E. (2009). A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 1-11.

Benton, A.L., Hamsher, K, D., Varney, N.R. and Spreen, O. (1983). *Contributions to Neuropsychological Assessment: A Clinical Manual*. New York: Oxford University Press.

Best, C., Moffat, V., Power, M., Owens, D., & Johnstone, E. (2008). The boundaries of the cognitive phenotype of autism: theory of mind, central coherence and ambiguous figure perception in young people with autistic traits. *Journal of Autism and Developmental Disorders, 38*, 840-847.

Boomsma, A., Van Lang, N., De Jonge, M., De Bildt, A., Van Engeland, H., & Minderaa, R. (2008). A new symptom model for autism cross-validated in an independent sample. *Journal of Child Psychology and Psychiatry, 49,* 809-816.

Boyd, B, Conroy, M., Mancil, G., Nakao, T., & Alter, P. (2007). Effects of circumscribed interests on the social behaviors of children with autism spectrum disorders. *Journal of*

Autism and Developmental Disorders, 37, 1550–1561.

Boyd, B., McBee, M., Holtzclaw, T., Baranek, G., & Bodfish, J. (2009). Relationships among repetitive behaviors, sensory features, and executive functions in high functioning autism. *Research in Autism Spectrum Disorders, 3,* 959–966.

Carroll, J., & Chiew, Y. (2006). Sex and discipline differences in empathizing, systemizing and autistic symptomology: evidence from a student population. *Journal of Autism and Developmental Disorders, 36,* 949-957.

Charman, T., (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society of London,* 358, 315-324.

Charman, T., Taylor, E., Drew, A., Cockeril, H., Brown, J., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry, 46,* 500–551.

Chen, G.M., Yoder, K.J., Ganzel, B.L., Goodwin, M.S., & Belmonte, M.K. (2012). Harnessing repetitive behaviors to engage attention and learning in a novel therapy for autism: An exploratory analysis. *Frontiers in Educational Psychology, 3,* 12.

Chen Y., Rodgers J, McConachie H. (2009). Restricted and repetitive behaviors, sensory processing and cognitive style in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 635–642.

Constantino, J., Gruber, C., Davis, S., Hayes, S., Passante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*, 719-726.

Cuccaro, M., Shao, Y., Grubber, J., Slifer, M., Wolpert, C. M., Donnelly, S., et al. (2003). Factor analysis of restricted and repetitive behaviors in autism using the Autism Diagnostic

Interview-R. *Child Psychiatry and Human Development, 34*, 3–17.

Dawson, G., Webb, S.J., Schellenberg, G.D., Dager, S, Friedman, S., Aylward, E., & Richards, T. (2002). Defining the broader phenotype of autism: genetic, brain, and behavioral perspectives. *Development and Psychopathology, 14,* 581-611.

Dworzynski, K., Ronald, A., Hayiou-Thomas, M., Rijsdijk, F., Happé, F., Bolton P., & Plomin, R. (2007). Aetiological relationship between language performance and autistic-like traits in childhood: a twin study. *International Journal of Language and Communication Disorders, 42,* 273-292.

Evans, D.W., Gray, F.L., & Leckman, J.F. (1999). The rituals, fears and phobias of young children: insights from development, psychopathology and neurobiology. *Child Psychiatry and Human Development 29,* 261-276.

Evers, K., Noens, I., Steyaert, J., & Wagemans, J. (2011). Combining strengths and weaknesses in visual perception of children with an autism spectrum disorder: perceptual matching of facial expressions. *Research in Autism Spectrum Disorders, 5,* 1327-1342.

Frazier, T., Youngstrom, E., Kubu, C., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders, 38,* 474-480.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition, 50,* 115-132.

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W., Goldberg, J., & Mahoney, W. (2007). Structure of the autism symptom phenotype: a proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry, 46,* 188-196.

Geschwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization: Biological mechanisms,*

associations, and pathology. Cambridge, MA:MIT Press.

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology, 18,* 591-617.

Gotham, K., Risi, S., Dawson, G., Tager-Flusberg, H., Joseph, R., Carter, A., et al. (2008). A replication of the Autism Diagnostic Observation Schedule (ADOS) revised algorithm. *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 642–651.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders, 37,* 613–627.

Happé, F., Briskman, J., & Frith, U. (2001). Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: I. experimental tests. *Journal of Child Psychology and Psychiatry, 42,* 299-307.

Happé, F., & Ronald, A. (2008). The 'fractionable triad': a review of evidence from behavioral, genetic, cognitive and neural research. *Neuropsychological Review, 18,* 287-304.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9,* 1218–1220.

Hoekstra, R., Bartels, M., Hudziak, J., Van Beiksterveldt, T., & Boomsma, D. (2008). Genetic and environmental co-variation between autistic traits and behavioral problems. *Twin Research and Human Genetics, 10,* 853-860.

Hoekstra, R., Bartels, M., Cath, D., & Boomsma, D. (2008). Factor Structure, Reliability and Criterion Validity of the Autism-Spectrum Quotient: A Study in Dutch Population and Patient Groups. *Journal of Autism and Developmental Disorders, 38,* 1555-1566.

Hurst, R., Mitchell, J., Kimbrel, N., Kwapil, T., & Nelson-Gray, R. (2007). Examination of the reliability and factor structure of the autism spectrum quotient (AQ) in a non-clinical sample. *Personality and Individual Differences, 43*, 1938-1949.

Jarrold, C., Butler, D. W., Cottington, E. M., & Jimenez, F. (2000). Linking theory of mind and central coherence in autism and in the general population. *Developmental Psychology 36,* 126–138.

Johnson, M.H., Halit, H., Grice, S.J., & Karmiloff-Smith, A. (2002). Neuroimaging of typical and atypical development: A perspective from multiple levels of analysis. *Development and Psychopathology, 14,* 521–536.

Kamp-Becker, I., Ghahreman, M., Smidt, J., & Remschmidt, H. (2009). Dimensional structure of the autism phenotype: relations between early development and current presentation. *Journal of Autism and Developmental Disorders, 39,* 557-571.

Kanner, L. (1943). Autistic disturbances of affective content. *Nervous Child, 2,* 217–250.

Kern, J., Trivedi, M., Grannemann, B., Garver, C., Johnson, D., Andrews, A., Savla, J., Mehta, J., & Schroeder, J. (2007). Sensory correlations in autism. *Autism, 11,* 123-134.

Kloosterman, P., Keefer, K., Kelley, E., Summerfeldt, L., Parker, J. (2011). Evaluation of the factor structure of the autism-spectrum quotient. *Personality and Individual Differences, 50,* 310-314.

Kolevzon, A., Smith, C., Schmeidler, J., Buxbaum, J., & Silverman, J. (2004). Familial symptom domains in monozygotic siblings with autism. *American Journal of Medical Genetics Part B, 129B,* 76-81.

Lam, K., Bodfish, J., & Piven, J. (2008). Evidence for three subtypes of repetitive behavior in autism that differ in familiality and association with other symptoms. *Journal of Child*

Psychology and Psychiatry, 49, 1193-1200.

Lord, C., Cook, E., Leventhal, B., & Amaral, D. (2000). Autism spectrum conditions. *Neuron, 28,* 355-363.

Magyer, C., & Pandolfi, V. (2007). Factor structure evaluation of the childhood autism rating scale. *Journal of Autism and Developmental Disorders, 37,* 1787-1794.

Mandy, W., & Skuse, D. (2008). Research review: what is the association between the social-communication element of autism and repetitive interests, behaviors and activities> *Journal of Child Psychology and Psychiatry, 49*, 795-808.

Matson, J., Boisjoli, J., Hess, J., & Wilkins, J. (2010). Factor structure and diagnostic fidelity of the Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT-Part 1). *Developmental Neurorehabilitation, 13,* 72–79.

McGovern, C., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry, 46,* 401-408.

Ozonoff, S., Pennington, B. R, & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32,* 1081-1105.

Papageorgiou, V., Georgiades, S., & Mavreas, V. (2008). Brief report: cross-cultural evidence for the heterogeneity of the restricted, repetitive behaviors and interests domain of autism: a Greek study. *Journal of Autism and Developmental Disorders, 38,* 558-561.

Pellicano, E., Maybery, M., and Durkin, K. (2005). Central coherence in typically developing preschoolers: does it cohere and does it relate to mindreading and executive control? *Journal of Child Psychology and Psychiatry, 46,* 533-547.

Piven, J. (1999). Genetic liability for autism: the behavioral expression in relatives. *International*

Review of Psychiatry, 11, 299-308.

Richler, J., Bishop, S. L., Kleinke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37,* 73–85.

Rogers, S. & Ozonoff, S. (2005). Annotation: what do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 46,* 1255–1268.

Ronald, A., Happé, F., Bolton, P., Butcher, L., Price, R. Wheelwright, S., Baron-Cohen, S., & Plomin, R. (2006a). Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 45,* 691-699.

Ronald, A., Happé, F., Price, T., Baron-Cohen, S., & Plomin, R. (2006b). Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *Journal of the American Academy of Child and Adolescent Psychiatry, 45,* 1206-1214.

Sacco, R., Curatolo, P., Manzi, B., Militerni, R., Bravaccio, C., Frolli, A., Lenti, C., Saccani, M., Elia, M., Reichelt, K.L., Pascucci, T., Puglisi-Allegra, S., Persico, A.M. (2010). Principal pathogenetic components and biological endophenotypes in autism spectrum disorders. *Autism Research, 3,* 237-252.

Silverman, J., Smith, C., Schmeidler, J., Hollander, E., Lawlor, B., Fitzgerald, M., Buxbaum, J., Delaney, K., Galvin, P., and the Autism Genetic Research Exchange Consortium. (2002). Symptom domains in autism and related conditions: evidence for familiality. *American Journal of Medical Genetics, 114,* 64-73.

Skuse, D., Mandy, W., Steer, C., Miller, L., Goodman, R., Lawrence, K., Emond, A., & Golding,

J. (2009). *Journal of the American Academy of Child and Adolescent Psychiatry, 48,* 128-137.

Snow, A., Lecavalier, L., & Houts, C. (2009). The structure of the autism diagnostic interview-revised: diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry, 50,* 734-742.

Spiker, D., Lotspeich, L., Dimiceli, S., Myers, R., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114,* 129-136.

Stewart, M., & Austin, E. (2009). The structure of the autism-spectrum quotient (AQ): evidence from a student sample in Scotland. *Personality and Individual Differences, 47,* 224-228.

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., Goldberg, J., & Tuff, L. (2006). Investigating the structure of the restricted, repetitive behaviors and interests domain of autism. *Journal of Child Psychology and Psychiatry, 47,* 582-590.

Szatmari, P., Merette, C., Bryson, S., Thivierge, J., Roy, M., Cayer, M., et al. (2002). Quantifying dimensions in autism: a factor-analytic study. *Journal of the American Academy of Child & Adolescent Psychiatry, 41,* 467–474.

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., & McGrath, L. (2003). A principal component analysis of the autism diagnostic interview-revised. *Journal of American Academy of Child Adolescent Psychiatry, 42,* 864–872.

Trope Y, Liberman N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117,* 440-463.

Turner, M. Annotation: repetitive behavior in autism: a review of psychological research. (1999).

Journal of Child Psychology and Psychiatry, 40, 839–849.

Valla, J.M., Ganzel, B.L., Yoder, K.J., Chen, G.M., Lyman, L.T., Sidari, A.P., Keller, A.E., Maendel, J.W., Perlman, J.E., Wong, S.K.L., & Belmonte, M.K. (2010). More than maths and mindreading: sex differences in empathizing-systemizing co-variance. *Autism Research, 3,* 174-184.

Van Lang, N., Boomsma, A., Sytema, S., de Bildt, A., Kraijer, D., Ketelaars, C., & Minderaa, R. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry, 47,* 37-44.

Voracek, M. & Dressler, S. (2006). Lack of correlation between digit ratio (2D4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemizing, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41,* 1481-1491.

Walter, E., Dassonville, P., and Bochsler, T. (2008). A specific autistic trait that modulates visuospatial illusion susceptibility. *Journal of Autism and Developmental Disorders, 39,* 339-349.

Watson, L., Patten, E., Baranek, G., Poe, M., Boyd, B., Freuler, A., & Lorenzi, J. (2011). Differential associations between sensory response patterns and language, social, and communication measures in children with autism or other developmental disabilities. *Journal of Speech, Language, and Hearing Research, 54,* 1562-1576.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103-128.

Witkin, H.A. (1950). Individual differences in ease of perception of embedded figures. *Journal of Personality, 19,* 1-15.

Witelson, S., & Nowakowski, R. (1991). Left out axons make men right: A hypothesis for the origin of handedness and functional asymmetry. *Neuropsychologia, 29,* 327–333.

Yoder, K.J., & Belmonte, M.K. (2011). Information processing and integration. In: *Autism Spectrum Disorders* (DG Amaral, G Dawson & DH Geschwind, eds.). New York: Oxford University Press, 2011, pp 1010-1027.