GENDER-SPECIFIC SENSITVITY TO TIME-DISCREPANT TASK CONDITIONS
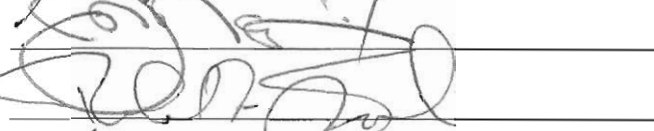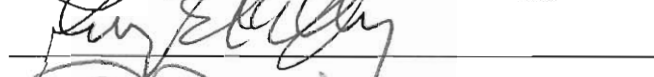OF REASONING DURING fMRI

by

Joshua M. Roberts
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Arts
Psychology

Committee:

_____ Director

_____

_____ Department Chairperson

_____ Dean, College of Humanities
and Social Sciences

Date: ___7. 22, 2011_____ Summer Semester 2011
George Mason University
Fairfax, VA

Gender-Specific Sensitivity to Time-Discrepant Task Conditions of Reasoning During fMRI

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Arts at George Mason University

By

Joshua M. Roberts
Bachelor of Arts
Richard Stockton College of NJ, 2001

Director: Layne Kalbfleisch, Professor
Department of Education

Summer Semester 2011
George Mason University
Fairfax, VA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

GENDER-SPECIFIC SENSITIVITY TO TIME-DISCREPANT TASK CONDITIONS
OF REASONING DURING FMRI

Joshua M. Roberts, M.A.

George Mason University, 2011

Thesis Director: M. Layne Kalbfleisch, M.Ed., Ph.D.

This study examined behavioral discrepancies and their neural correlates between males
and females on a color relational complexity (CRC) matrix reasoning task administered
in two separate functional magnetic resonance imaging (fMRI) studies. Using 42 age-
matched, CRC-trained participants, behavioral results revealed significant performance
differences by gender, indicating that males performed more accurately, but less rapidly
on CRC trials than females. These gender differences were also expressed in the neural
activation maps indicating conserved right lateralized activity for males and a robust
frontal and limbic associated activity for females, attributed to cognitive load associated
with stereotype threat. Consideration of timing parameters with a 'timed' versus a 'self-
paced' design elucidated further differentiation of the neural activation supporting color
relational complexity, with the self-paced condition supported by posterior visuo-spatial
regions and the timed condition supported by a sparse cingulate activation.

INTRODUCTION

Sex differences in task performance, for example spatial attention (Bell, Willson, Wilman, Dave, & Silverstone, 2006), mental rotation (Thomsen, et al., 2000), or Tower of London (Unterrainer, et al., 2005), and their corresponding neural activations have been well documented albeit with varying conclusions (Voyer, Voyer, & Bryden, 1995). Most of the relevant literature, however, uses three-dimensional spatial tasks to measure non-verbal reasoning, problem solving, and perception (Weiss, et al., 2003). The present study seeks to extend this cognitive measurement by employing the use of a two-dimensional matrix reasoning task. This measure, a common analogue measure of aptitude assessment employed in tests such as the Raven's Progressive Matrices (1947) and the Naglieri Nonverbal Ability Test (1997), has implications for documenting this form of visual-spatial cognition as it bears on the assessment of human ability and aptitude. Furthermore, it will supplement recent findings from developmental cognitive neuroscience on sex differences in a spatial line orientation task that also document visual-spatial cognition in a two-dimensional perceptual space (Clements-Stephens, Rimrodt, & Cutting, 2009).

The matrix reasoning measure explores the phenomenon of relational complexity (Christoff, et al., 2001). Relational complexity (RC) refers to the number of relations (or components of a stimulus) that need to be simultaneously considered to determine an accurate solution (Kroger, et al., 2002). Our lab recently characterized the impact of color on the concept of relational complexity, thus, gender differences will be explored within conditions on this task (Kalbfleisch, et al., in submission). The task stimuli in this study are color matrix reasoning problems constructed to examine relational complexity between 1-relational and 2-relational levels (see Figure 1).

Past research on gender differences in spatial ability has largely relied on three dimensional object rotation tasks (Voyer, et al, 1995), which have been shown to produce very robust neural maps and elicit neural and behavioral gender distinctions. Other studies using non-verbal reasoning tasks, have utilized 2-Dimensional line drawings (Kroger, et al., 2002; Gur, et al., 2000; Clements-Stevens, et al., 2009) in an effort to minimize drastic variance in performance and strip down the neural network of visuospatial processing to its core components, getting rid of any extraneous cognitive demands beyond the scope of interest. However, since we experience the world in 3D and not 2D, this transition to 2-dimensional models results in the loss of some ecological validity. In an effort to represent a closer approximation to real-world perception in a 2D matrix, we added color to our matrix reasoning stimuli.

In addition to this, we employed the use of event-related fMRI experimental design to create further ecological validity in our task design. There are two types of functional magnetic resonance imaging (fMRI) designs, block design and event-related design.

Block designs rely on successive presentation of the same trial type for a discretely defined period of time. Event-related design, on the other hand, allows for multiple trial types to be presented successively in rapid and random intervals. This results in a design that is more ecologically valid than that obtained from block design Event-related design in fMRI has the advantage of keeping the subject on-task to prevent incidental cognitive processing, and reduces the likelihood that they will anticipate or prepare for the next stimuli or condition (Donaldson and Buckner, 2001). In an effort to produce a task more fitting to actual cognition we employed the use of event-related design.

The involvement of the prefrontal cortex (PFC) in relational complexity processing is well chronicled, suggesting not only that proper functioning of the PFC is essential to multiple relation synthesis (Waltz, et al., 2002), but that PFC activation increases with increasing complexity (Christoff et al., 2001). Christoff and colleagues more specifically characterized this PFC activation in the bilateral rostro-lateral PFC (RLPFC), noting a switch to left hemispheric activity as relational complexity increased, and indicated the involvement of the right dorsolateral PFC (DLPFC) in spatial processing and working memory. Kroger and colleagues (2002) also implicated the bilateral DLPFC along with the parietal region in fluid reasoning and working memory processes. This parietal activation is complemented by Prabhakran, Rypma, and Gabrieli (2001) showing increased activation in this area during figural processing tasks. As relational complexity increased, Kroger and colleagues (2002) noted greater activation in the anterior left PFC. These regions are thought to provide the neural basis required for accurate performance during matrix reasoning (Kalbfleisch, 2004).

Generally it has been proposed that men perform spatial processing tasks differently than women due to different strategies. Men are thought to implement a more efficient bottom-up approach or gestalt method, whereas women rely on top-down processing/serial reasoning (Butler, et al., 2006; Thomsen, et al., 2000). Neural activations noted during these spatial tasks vary but some commonalities do emerge. Males have shown greater parietal activation overall (Weiss, et al., 2003; Thomsen, et al., 2000; Jordan, Wustenburg, Heinze, Peters, & Jancke, 2002; Hugdahl, Thomsen, & Ersland, 2006; Gur, et al., 2000) and women greater frontal lobe activation (Butler, et al., 2006; Weiss, et al., 2003; Thomsen, et al., 2000; Hugdahl, et al., 2006; Gill & O'Boyle, 1997). In addition to these general findings, other idiosyncratic gender differences in spatial processing tasks indicate that males engage the bilateral postcentral gyrus, precentral gyrus, precuneus and basal ganglia (Butler, et al., 2006), the right parieto-occipital sulcus, the left intraparietal sulcus and left superior lobule (Jordan, et al., 2002), right lingual gyrus and the cerebellum (Clements-Stevens, et al., 2009), and general bilateral engagement (Clements, et al., 2006) Females on the same task engage the bilateral dorsomedial PFC, right inferior temporal gyrus, the left inferior occipital lobe (Butler, et al., 2006), and present bilateral activation in intrapareital sulcus, superior and inferior parietal lobule, inferior temporal gyrus, and premotor areas (Jordan, et al., 2002), and general right lateralized activation (Clements, et al., 2006).

Table 1: Summary of neural activation as noted in previous visuospatial imaging research (as cited)		
Male	Common Activation	Females
• Postcentral gyrus • Precuneus • Basal ganglia • Parieto-occipital • Lingual gyrus • Cerebellum • General bilateral engagement	• Postcentral gyrus • Precuneus • Intraparietal • Superior parietal • Premotor	• Dorsomedial PFC • Inferior temporal • Inferior occipital • Inferior parietal • Inferior temporal • General right lateralized activation

Despite these findings, other researchers revealed no gender specific activations for their visuospatial task and concluded that any changes in neuronal activity were performance based, not gender based (Unterrainer, et al., 2005). Bell and colleagues, in a study based on fMRI (2006) also revealed no significant gender specific activations in a cognitive spatial attention task.

Complementary to this, some research indicates less distinct gender divergence on behavioral performance as well. Lack of gender differentiation in behavioral performance has been recorded in the Naglieri Non-Verbal Ability Test (NNAT) (Rohajn & Naglieri, 2005), mental rotation tasks (Hugdahl, et al., 2006; Christova, Lewis, Tagaris, Ugerbil, & Georgopoulos, 2008), and math skills (Fenema & Sherman, 1977). Indeed, based on a meta-analysis of spatial abilities, Voyer and colleagues (1995) report that cognitive sex differences may be getting smaller, possibly due to a departure from gender-based role expectations and subsequent rearing and educational tendencies. It should be noted,

however, that significant sex differences were found for several tests. Specifically males performed more favorably with spatial tasks, notably on the mental rotation task.

For this reason we examined potential neural and behavioral distinction between males and females during functional magnetic resonance imaging (fMRI) on the color relational complexity matrix reasoning task (CRC). We hypothesize that as relational complexity increases subjects will require greater cognitive effort and attention. Therefore we expect to see slower reaction times and decreased accuracy in the 2-relational complexity trials as compared to the 1-relational complexity trials. This increase in complexity and assumed cognitive demand will therefore require recruitment of the anterior left prefrontal cortex (Kroger, et al., 2002) and rostrolateral/dorsolateral PFC (Christoff, et al., 2001). Additionally, based on activations detected in previously documented line-orientation tasks (Gur, et al., 2007), we hypothesize the emergence of distinct neural networks, with males presenting posterior bilateral/parietal activation and females presenting primarily right hemispheric activation.

No previous studies examine performance differences based on gender when examining this perceptual process affiliated with matrix reasoning. This project aims to provide (1) provide a better understanding of gender differences in cognition by characterizing the behavioral and neural discrepancies between male and female subjects during matrix reasoning, (2) explore the possibility of gender differences in relational complexity, and (3) determine the impact of timing parameters on performance between genders.

METHODS

Subjects

Behavioral data from 80 right handed, non-colorblind participants will be analyzed for this experiment (40 male/40 female). 44 subjects were derived from Study A and 36 from Study B. Participants were matched for age resulting in a mean ages of 24.05 ($\pm$6.57) for males and 24.58 ($\pm$6.70) for females.

Cognitive Task

To solve each problem, participants consider the relationships among the 8 provided matrix blocks in the trial and select the correct answer from the four choices provided. The task involves two different types of relational complexity (RC): 1-relational and 2-relational. The 1-relational problems only have one single change either between columns or between rows of the matrix reasoning trial and they required processing for single relation (Fig. 1-A). 2-relational problems have at least two changes, both between columns and between rows and they require simultaneous integration of the processing for two or more relations (Fig. 1-B). Items in the color relational complexity are colored light blue, a color universally perceived by individuals with typical color perception as well as those with most common forms of color blindness (Deeb, 2004).
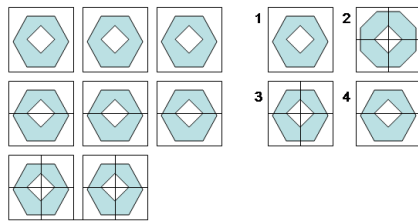
A. 1-relational complexity item
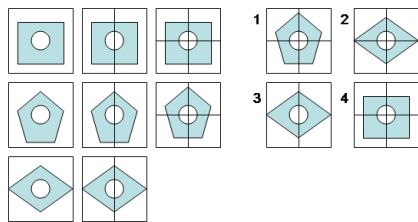
B. 2-relational complexity item

Figure 1. Color relational complexity

Subjects used their right hand index, middle, ring, and small fingers to press one of four buttons corresponding to their choice. All subjects were trained outside of the scanner to familiarize themselves with the task and within the scanner to orient them to the use of the button boxes prior to their functional scan. Problems were presented from a computer and back-projected through a wave-guide onto a screen located in the back of the MRI bore. Visual images were viewed from a mirror mounted on the head coil above the participant's head.

Behavioral Procedure

Study A

Relational complexity trials included a mixture of color, no-color, and visual contrast configured stimuli at three levels of complexity (0, 1, and 2+). Subjects were provided

the opportunity to go at their own pace and were informed that it was "more important to get the right answer than to finish quickly." Runs consisted of 117 unique relational complexity stimuli presented randomly with an inter-trial interval (ITI) of 3 seconds. Subjects completed a total of 3 runs per session. Sessions were counterbalanced across subjects. Reaction time was defined as the duration between onset of the stimuli and the response given.

Study B

Relational complexity trials were randomly presented using the color condition only at two levels of complexity (1 and 2+). Subjects were presented with randomized relational complexity stimuli and provided 3 seconds of study time. Following this study time they were cued to respond through presentation of a visual (green cross) and audible (tone) signal. Subjects were allotted 2895ms to choose their answer before being presented with another trial. Within each run the presentation of the stimuli and ITI was arranged using Optseq (MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging) to ensure the optimum randomization of stimuli presentation. The ITI varied according to the repetition time (RT) of the scan at 0, 3, 6, 9, or 12 seconds. Each run consisted of 60 unique relational complexity stimuli. The stimuli were identical to those used in Study A and the maximum response time imposed in Study B were informed by the performance on the self-paced version in Study A assuring that the timed condition would impose a strict, but manageable, pressured response. Reaction time was defined as the duration between onset of cued response (green cross and tone) and time of response (button press).

MRI Data Acquisition

Structural and functional MRI data were acquired using a 3.0-tesla Siemens Allegra head-only scanner (Siemens Medical Solutions, USA) located at the Krasnow Institute for Advanced Study at George Mason University. A CP TX/R head coil single-channel was used during data collection and head motion was restricted using memory foam inserts in the head coil. For each functional run, a standard BOLD (blood-oxygenation-level-dependent) gradient-echo echo-planner imaging (EPI) pulse sequence (TR = 3000 ms, TE = 30 ms, flip angle = 70º, FOV = 192 × 192 mm^2, 64 × 64 voxels) was used. The first three volumes acquired at the beginning of each session were discarded in order to account for signal stabilization. The start of each run was manually synchronized with the scanner. Structural images were collected as high-resolution T1-weighted images using a MPRAGE (magnetization-prepared rapid-acquisition gradient-echo) sequence with following parameters: (TR = 2300 ms, TE = 2.7 ms, flip angle = 12º, matrix size of 256 × 256) and additional structural images with following parameters: spin-echo axial-oblique T1-weighted structural scans of the whole brain (coplanar with functional images, 48 slices, repetition time (TR) = 200ms, echo time (TE) = 3.6 ms, field of view (FOV) = 192 × 192 mm^2; slice thickness = 3 mm, flip angle = 75º; matrix size = 205 × 256).

Preprocessing

Image reconstruction was performed offline. The conversion of raw data completed using MRIConvert (University of Oregon Lewis Center for Neuroimaging). Data processing and analysis were carried out using the Statistical Parametric Mapping

software package SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK). Pre-processing included slice-timing, reorientation, realignment using INRIAlign (Freire & Mangin, 2001), normalization to SPM EPI template, and spatial smoothing with 9 mm full width at half maximum (FWHM) isotropic Gaussian kernel. The movement inclusion criteria (less than 2 mm translational move and less than 2 degree rotational move) was met for all datasets. A high-pass filter with cut-off period of 128 s was used to remove low-frequency drifts unrelated to the experimental paradigm.

Statistical Analysis

Since we are only interested in gender differences present in color relational complexity trials at the 1- and 2(+)- relational complexity level, these trials were examined from 3 runs of the self-paced task in Study A. for a total of 3,432 trials. Data derived from the color relational complexity trials timed task condition in Study B, were selected from two runs and totaled 4,320 number of trials. Subject performance was determined through use of analyses of variance on reaction times for correctly solved trials and on accuracy rates. Mean values were calculated for both dependent variables at each level of complexity by gender.

Behavioral results are presented for all relevant subjects meeting criteria in the study. An additional subset of these behavioral results is also presented, comprised of only those subjects whose imaging results were included in the final functional MRI analysis.

Based on the differing task designs used, preliminary data analysis was conducted on each study independently prior to combining them into one dataset. Furthermore, based

on the variance and range in reaction times as imposed by the two task designs, RT's were independently transformed into z-scores. This transformation served as an additional comparative control to ensure validity of results and to create a high level of conformity between the data sets

Study A - All subjects:

Results were derived from 44 age-matched participants with 22 males (mean age = 26.2±6.82) and 22 females (mean age = 26.0±8.12).

Overall accuracy was high with a mean of 94.01% (±.24) on all trials. Results of an ANOVA on accuracy indicated no significant main effects for gender, $F(1, 3220) = 1.88$, $p = .17$, or complexity, $F(1, 3220) = 1.28$, $p = .26$. Based on reaction times however, significant main effects were revealed for both gender, $F(1, 3027) = 81.39$, and complexity, $F(1, 3027) = 151.25$, $p = .00$, indicating faster reaction times for females ($M = 4426$ ms) than for males ($M = 5441$ ms) overall and, as expected, slower reaction times on 2(+)-relational complexity trials ($M = 5607$ ms) than 1-relational complexity trials ($M = 4231$ ms) (see Table 1). Interaction effects were not significant ($p = .48$). Using the z-score transformed reaction time (RT) values yielded identical significant contrasts. Specifically, there were main effects for gender, $F(1, 3741) = 62.19$, and for complexity, $F(1, 3741) = 272.6$, $p = .00$.

Table 2: Mean accuracy and reaction times for all subjects by condition for Study A.

CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	94.6% (22.6)	93.6% (24.5)	95.6% (20.6)
	RT	4.23 s (2.75)	4.72 s (3.12)	3.79 s (2.27)
Color 2+	ACC	93.4% (24.8)	93.5% (24.7)	93.4% (24.9)
	RT	5.61 s (3.44)	6.17 s (4.17)	5.08 s (2.47)
Overall	ACC	94.0% (23.7)	93.5% (24.6)	94.5% (22.9)
	RT	4.91 s (3.19)	5.44 s (3.75)	4.42 s (2.45)

Effect size was also determined for reaction times with a point-biserial Pearson correlation, indicating significant effect sizes ($p < .01$) for gender, $r_{pb}(3029) = .16$, and complexity, $r_{pb}(3029) = .22$. Interaction effects were not significant ($p = 0.39$). RT effect sizes based on the transformed z-score values revealed significant effects sizes ($p < .01$) for complexity, $r_{pb}(2904) = .29$, and gender, $r_{pb}(2904) = .14$.

Study A - Scanned subjects only:

Results were derived from 22 age-matched participants with 11 males (mean age = 26.0±8.22) and 11 females (mean age = 24.9±8.23).

Overall accuracy was high with a mean of 96.27% (±.19) on all trials. Results of an ANOVA on accuracy indicated no significant main effects for complexity, $F(1, 1686) = 0.81, p = .37$. There were significant main effects for gender, $F(1, 1686) = 11.23, p = .00$, indicating greater accuracy in males ($M = 0.98$) than in females ($M = 0.95$).

Interaction effects for accuracy were not significant, $p = 0.71$. Based on reaction times, significant main effects were revealed for both gender, $F(1, 1686) = 18.43$, and complexity, $F(1, 1686) = 168.57$, $p = .00$, indicating significantly faster reaction times for females ($M = 4223$ ms) than for males ($M = 4677$ ms) overall and slower reaction times on 2(+)-relational complexity trials ($M = 5133$ ms) than 1-relational complexity trials ($M = 3761$ ms) (See Table 2). Interaction effects were not significant ($p = .68$). Using the z-score transformed RT values yielded identical significant contrasts. Specifically, there were main effects for gender, $F(1, 1589) = 24.91$, and for complexity, $F(1, 1589) = 193.2$, $p = .00$. Interaction effects were again not significant ($p = .42$).

Table 3: Mean accuracy and reaction times for scanned subjects by condition for Study A.

CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	96.7% (17.9)	98.1% (13.8)	95.3% (21.1)
	RT	3.76 s (1.89)	4.01 s (1.89)	3.52 s (1.86)
Color 2+	ACC	95.9% (19.9)	97.6% (15.3)	94.2% (23.5)
	RT	5.13 s (2.44)	5.34 s (2.47)	4.93 s (2.39)
Overall	ACC	96.3% (19.0)	97.8% (14.6)	94.8% (23.5)
	RT	4.47 s (2.29)	4.68 s (2.30)	4.22 s (2.26)

Effect size was also determined for reaction times indicating significant effect sizes ($p < .01$) for gender, $r_{pb}(1688) = .10$, and complexity, $r_{pb}(1688) = .30$. Effect size for accuracy was also significant for gender at the .01 level $r_{pb}(1688) = .08$. RT effect sizes

based on the transformed z-score values also revealed significant effects sizes ($p < .01$) for complexity, $r_{pb}(1591) = .33$, and gender, $r_{pb}(1591) = .12$.

Study B - All subjects:

Results were derived from 36 age-matched participants with 18 males (mean age = 21.4±5.29) and 18 females (mean age = 22.8±3.92).

Overall accuracy was high, with a mean of 91% (±.29) on all trials. Results of an ANOVA on accuracy indicated significant main effects for gender, $F(1, 4316) = 13.58$, and complexity, $F(1, 4316) = 72.15$, $p = .00$, indicating lower overall accuracy for females ($M = .89$) than males ($M = .93$) and greater accuracy on 1-relational complexity trials ($M = .95$) than on 2(+)-relational complexity trials ($M = .87$). Interaction effect of gender and complexity also reached significance, $F(1, 4316) = 22.6, p = .00$, indicating that females were more accurate than males at the 1-relational complexity level ($M = .95$, $M = .94$ respectively), but less accurate at the 2(+)-relational complexity level ($M = .84$, $M = .91$). Based on reaction times, there were significant main effects for gender, $F(1, 3921) = 35.38$, and complexity, $F(1, 3921) = 102.75$, $p = .00$, indicating faster reaction times for females ($M = 842$ ms) than for males ($M = 924$ ms) overall, and slower reaction times on 2(+)-relational complexity trials ($M = 955$ ms) than on 1-relational complexity trials ($M = 818$ ms) (See Table 3). Interaction effect of reaction times for gender and complexity were not significant ($p = .95$). Using the z-score transformed RT values yielded identical significant contrasts. Specifically, there were main effects for gender,

$F(1, 3741) = 25.29$, and for complexity, $F(1, 3741) = 84.34$, $p = .00$. Interaction effects were again not significant ($p = 0.22$).

Table 4: Mean accuracy and reaction times for all subjects by condition for Study B.				
CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	94.5% (22.7)	94.1% (23.6)	95.0% (21.8)
	RT	818 ms (391)	858 ms (415)	778 ms (362)
Color 2+	ACC	87.2% (33.4)	90.8% (28.9)	83.5% (37.1)
	RT	955 ms (451)	993 ms (487)	914 ms (404)
Overall	ACC	90.9% (28.8)	92.5% (26.4)	89.3% (31.0)
	RT	884 ms (426)	924 ms (456)	842 ms (388)

Effect size was also determined for reaction times indicating significant effect sizes (*p <.01)* for gender, $r_{pb}(3923) = .10$, and complexity $r_{pb}(3923) = .16$. Effect sizes for ACC also indicated significance at the .01 level for gender $r_{pb}(4318) = .06$, and complexity, $r_{pb}(4318) = .13$. RT effect sizes based on the transformed z-score values revealed significant effects sizes (*p < .01)* for complexity, $r_{pb}(3743) = .15$, and gender, $r_{pb}(3743) = .09$.

Study B - Scanned subjects only:

Results were derived from 20 age-matched participants with 10 males (mean age = 21.1±4.07) and 10 females (mean age = 22.5±3.66).

Overall accuracy was high, with a mean of 94.5% (±.23) on all trials. Results of an ANOVA on accuracy indicated significant main effects for gender, $F(1, 2396) = 15.79$,

and complexity, $F(1, 2396) = 23.78$, $p = .00$), indicating lower overall accuracy for females ($M = .93$) than males ($M = .96$) and greater accuracy on 1-relational complexity trials ($M = .97$) than on 2(+)-relational complexity trials ($M = .92$). Interaction effect of gender and complexity also reached significance, $F(1, 2396) = 5.51$, $p = .02$, indicating that although females were less accurate than males at both levels of complexity, the discrepancy was greater in the 2(+)-relational level. Based on reaction times, there were again significant main effects for gender, $F(1, 2264) = 159.68$, and complexity, $F(1, 2264) = 51.86$, $p = .00$, indicating faster reaction times for females ($M = 756$ ms) than for males ($M = 975$ ms) overall and slower reaction times on 2(+)-relational complexity trials ($M = 933$ ms) than on 1-relational complexity trials ($M = 806$ ms) (See Table 4). Interaction effect of reaction time for gender and complexity were not significant ($p = .94$). Using the z-score transformed RT values yielded identical significant contrasts. Specifically, there were main effects for gender, $F(1, 2167) = 148.32$, and for complexity, $F(1, 2167) = 41.97$, $p = .00$. Interaction effects were not significant ($p = 0.16$).

Table 5: Mean accuracy and reaction times for scanned subjects by condition for Study B.

CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	96.8% (17.7)	97.5% (15.6)	96.0% (19.6)
	RT	806 ms (393)	914 ms (419)	696 ms (332)
Color 2+	ACC	92.3% (26.7)	95.2% (21.5)	89.3% (30.9)
	RT	933 ms (453)	1037 ms (485)	821 ms (386)
Overall	ACC	94.5% (22.8)	96.3% (18.8)	92.7% (26.1)
	RT	868 ms (428)	975 ms (457)	756 ms (364)

Effect size was also determined for reaction times indicating significant effect sizes ($p < .01$) for gender, $r_{pb}(2266) = .26$, and complexity, $r_{pb}(2266) = .15$. Effect sizes for ACC also indicated significance at the .01 level for gender, $r_{pb}(2398) = .08$ and complexity, $r_{pb}(2398) = .10$. RT effect sizes based on the transformed z-score values revealed significant effects sizes ($p < .01$) for complexity, $r_{pb}(2169) = .13$, and gender, $r_{pb}(2169) = .25$.

Examining both of these datasets (and their associated subsets) individually indicates similar patterns of subject performance. Commonality across the two studies indicates the following: 1) males are significantly more accurate than females, 2) females are significantly faster in accurate problem-solving than males, and 3) 1-relational complexity trials were solved significantly faster than 2(+)-relational trials. This comparison therefore indicates a high level of consistency in task performance across

studies (Study A vs. Study B), as well as the representative sample used for the fMRI

analysis (All Subjects vs. Scanned Subjects).

Table 6: Behavioral results on accuracy and reaction time based on task parameters
(Study A vs. Study B), complexity (Color 1 vs. Color 2+), and gender.

		Self-Paced	Significance	Timed	Significance
Accuracy	Males	97.8% (14.6)	$F = 11.23$	96.3% (18.8)	$F = 15.79$
	Females	94.8% (23.5)		92.7% (26.1)	
	Color 1	96.7% (17.9)	Non-significant	96.8% (17.7)	$F = 23.78$
	Color 2+	93.9% (19.9)		92.3% (26.7)	
Reaction Times	Males	4.68s (2.30)	$F = 18.43$	975ms (457)	$F = 159.68$
	Females	4.22s (2.26)		756ms (364)	
	Color 1	3.76s (1.89)	$F = 168.57$	806ms (393)	$F = 51.86$
	Color 2+	5.13s (2.44)		933ms (453)	

Only one significant effect was measured between studies. Study B indicated

significant main effects for accuracy, as well as an interaction effect for accuracy and

gender whereas Study A did not. We attribute this to the self-paced design of Study A

which provided ample time for consistent and accurate problem solving. This finding is

not detrimental to the strategy for merging the two sets of data as we did not expect

significant gender differences related to accuracy in Study A.

Combined Study:

All subjects

The merged data set is comprised of 80 age-matched participants with 40 males

(mean age = 24.1±6.57) and 40 females (mean age = 24.6±6.70).

Subject accuracy was assessed at a mean of 92.21% (±0.27) on all trials. Results of the ANOVA on accuracy indicated there were significant main effects for gender, $F(1, 7540) = 5.02, p = .03$, and complexity, $F(1, 7540) = 58.11, p = .00$, indicating lower overall accuracy for females ($M = .92$) than males ($M = .93$) and greater accuracy on 1-relational complexity trials ($M = .95$) than on 2(+)-relational complexity trials ($M = .90$). Interaction effects of complexity and gender also reached significance, $F(1, 7540) = 19.99, p = .00$, whereby females performed more accurately than males at the 1-relational complexity level ($M = .95, M = .94$ respectively), but males performed more accurately than females at the 2(+)-relational complexity level ($M = .92, M = .88$ respectively)

Based on reaction times, there were significant main effects for gender, $F(1, 6952) = 28.17$, and complexity, $F(1, 6952) = 113.83, p = .00$, indicating faster reaction times for females ($M = 2452$ ms) than males ($M = 2831$ ms) overall and slower reaction times on 2(+)-relational complexity trials ($M = 3023$ ms) than 1-relational complexity trials ($M = 2277$ ms) for all subjects (See Table 5). Interaction effects for reaction time and complexity were not significant ($p = .86$). Using the z-score transformed RT values yielded identical significant contrasts. Specifically, there were main effects for gender, $F(1, 6647) = 72.90$, and for complexity, $F(1, 6647) = 282.98, p = .00$. Interaction effects were not significant ($p = 0.11$).

Table 7:
Mean accuracy and reaction times for all subjects by condition for Combined Study.

CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	94.6% (22.7)	93.9% (24.0)	95.2% (21.3)
	RT	2.28s (2.48)	2.47s (2.79)	2.09s (2.13)
Color 2+	ACC	90.0% (30.2)	91.9% (27.2)	87.8% (32.7)
	RT	3.02s (3.27)	3.20s (3.75)	2.84s (2.69)
Overall	ACC	92.2% (26.8)	92.9% (25.7)	91.5% (27.9)
	RT	2.64s (2.92)	2.83s (3.32)	2.09s (2.13)

Effect size was determined through a point-biserial Pearson correlation. Significant effect size of measured accuracy was found for gender, r_{pb} (7542) = .03, and complexity, r_{pb} (7542) = .09, (p < .01). Significant effect size of measured reaction time was also found for gender, r_{pb} (6954) =.07, and complexity r_{pb} (6954) =.13, (p < .01). Using the transformed z-score RT's also yielded significant correlation at the .01 level for complexity, r_{pb} (6649) =.20, and gender, r_{pb} (6649) =.11.

fMRI Analysis

General linear model (GLM) is the classic statistical method for fMRI analysis (Penny, Holmes, & Friston, 2003). A GLM is applied here to the time course of fMRI activation in the brain to estimate condition effects at the voxel level. In fMRI analysis, the brain is parsed into 3x3 millimeter voxels as a unit of measure similar to the way 2-dimensional pixels comprise digital images. In the first-level (single subject) analysis, the

response function will be modeled as an event-related function matched to the participant's moment of response convolved with a canonical hemodynamic response function (Friston, et al., 1994). A linear contrast of estimated regression coefficients is used to compare the specific effects of each condition. The fMRI data from each participant are used to generate statistical contrasts for brain activation related to specific task conditions. Data is evaluated at the single subject level to disqualify data that include activation outside of neural tissue, an indication of motion and/or physiological artifact. In addition, each trial is verified for behavioral accuracy for inclusion in the analysis. Based on the fidelity of the single subject data, the resulting statistical parametric maps yield a t-statistic of each voxel value for each contrast in first level analysis. These data are then entered into the second-level (group) analysis. A random effects model is used to account for both scan-to-scan and subject-to-subject variability. In the second level, one-sample t-tests are applied to the first-level statistical parametric maps to correlate behavioral parameters with the blood-oxygen dependent (BOLD) signal, the cardinal measure of fMRI. Specific contrasts include color relational complexity (CRC) at level 2+ versus color relational complexity at level 1, CRC at level 2+ versus CRC at level 1 for males, CRC at level 2+ versus CRC at level 1 for females, and CRC at level 2+ for males versus CRC at level 2+ for females. These contrasts were additionally interrogated to parse out activation by task condition (Study A vs. Study B). A secondary analysis was conducted applied to the same subjects and contrasts to covary for participant age in order to rule out any potential age-related developmental discrepancy. Applying the rules of classic fMRI analysis, the resulting statistical maps are statistically determined at an

uncorrected threshold of p<0.001 and indicated by a cluster (k) extent size of more than 10 voxels to identify the areas of neural activation related to each contrast. Peak activations are corrected for multiple comparisons using a false discovery rate (FDR) of p <. 05 as indicated.

RESULTS

Behavioral Performance

Those individuals whose post-processed fMRI data exceeded the limits of the motion correction software (>2mm) and those with excessive artifact were excluded from the image analysis results. Additionally, only those runs and/or subjects meeting the predetermined behavioral criteria of 85% accuracy or above (per run) were included in the data analysis. Following gender and age-matching, subjects consisted of 42 age-matched participants with 21 males (mean age = 23.7±6.89) and 21 females (mean age = 23.8±6.43).

Overall, subject accuracy is measured by a mean of 95.23% (±0.21) on all trials. Results of the ANOVA on accuracy yield significant main effects for gender, $F(1, 4086)$ = 26.51, and complexity, $F(1, 4086) = 20.15$, $p = .00$, indicating lower overall accuracy for females ($M = .94$) than for males ($M = .97$) and greater accuracy on 1-relational complexity trials ($M = .97$) than on 2(+)-relational complexity trials ($M = .94$). Interaction effects of complexity and gender also reach significance, $F(1, 4086) = 4.46$, $p = .04$. This suggests that although females performed significantly worse at both levels of complexity, there was greater discrepancy in accuracy at the 2(+)-relational level (See Figure 2).
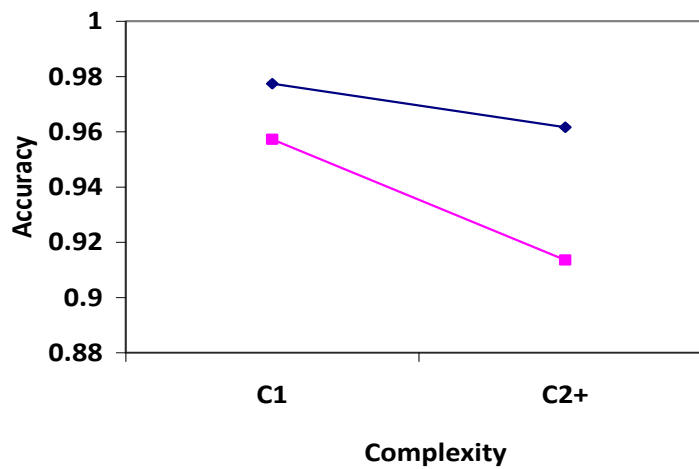
Figure 2: Performance on color relational complexity by complexity and gender for scanned subjects.

A reaction time (RT) analysis yields significant main effects for gender, $F(1, 3891) = 13.80$, and complexity, $F(1, 3891) = 85.54$, $p = .00$, indicating faster reaction times for females ($M = 2225$ ms) than males ($M = 2505$ ms) overall and slower reaction times on 2(+)-relational complexity trials ($M = 2713$ ms) than 1-relational complexity trials ($M = 2030$ ms) for all subjects. Interaction effects for reaction time and gender were not significant ($p = .44$) (See Figure 3 and Table 6). These results are mimicked by the z-score transformed RT's which reveals significant main effects for gender, $F(1, 3760) = 166.72$, and complexity, $F(1, 3760) = 160.86$, $p = .00$.
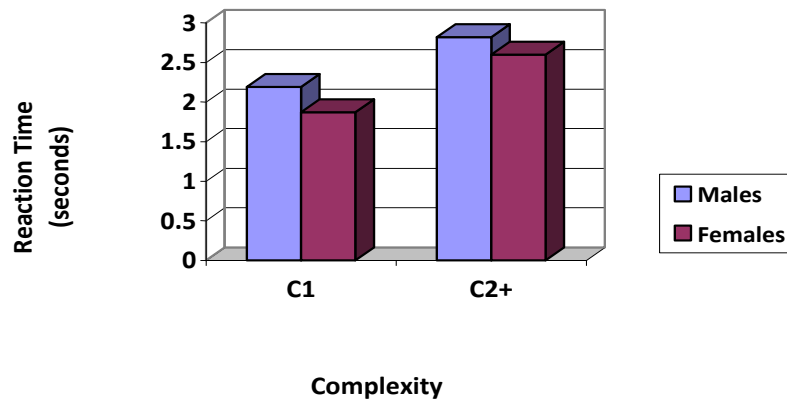
Figure 3: Performance on color relational complexity by complexity and gender for scanned subjects.

Table 8:
Mean accuracy and reaction times for scanned subjects by condition for Combined Study.

CONDITION		OVERALL	MALE	FEMALE
Color 1	ACC	96.7% (17.8)	97.0% (17.2)	93.5% (24.6)
	RT	2.03s (1.93)	2.19s (1.98)	1.87s (1.87)
Color 2+	ACC	93.7% (24.2)	96.2% (19.2)	91.4% (28.1)
	RT	2.71s (2.63)	2.82s (2.66)	2.60s (2.61)
Overall	ACC	95.2% (21.3)	97.0% (17.2)	93.5% (24.6)
	RT	2.37s (2.33)	2.50s (2.36)	2.23s (2.29)

Effect size was determined through a point-biserial Pearson correlation. Significant effect size of measured accuracy was found for gender, $r_{pb}(4088) = .08$ and complexity, $r_{pb}(4088) = .07$, p < .01. Significant effect size of measured reaction time was also found for gender, $r_{pb}(3893) = .06$, and complexity, $r_{pb}(3893) = .15$, $p < .01$. Using the

transformed z-score RT's also yielded significant correlation effects with complexity, $r_{pb}(3762) = .20$, and gender, $r_{pb}(3762) = .20$, $p < .01$.

Neuroimaging Results

BOLD activation maps resulting from the GLM analyses are outlined in Tables 7-10 and illustrated in Figures 4-7. MNI coordinates derived from SPM5 output were translated to Talairach coordinates and subsequently identified via the Talairach Client Deamon (Lancaster, et al., 1997; Lancaster, et al., 2000). Overall, maps comprised of both genders (N = 42) contrasting 2(+)-relational vs. 1-relational trials indicate activation in the right middle frontal gyrus (BA 9) and right precentral gyrus at a cluster level correction of p < .05 (See Figure 4). The 2-relational vs. 1-relational contrast for Study A (N = 22) revealed active clusters (FDR correction <.05) in the right superior parietal lobule and bilateral precuneus (BA 7), right inferior frontal gyrus (BA 9), right middle frontal gyrus (BA 6), left angular gyrus (BA 39), right middle and left inferior occipital gyrus, and left cuneus (BA 18), as well as, left cerebellar declive. The 2(+)-relational vs. 1-relational contrast for Study B (N = 20) reveals cluster corrected (p < .05) activation in the left cingulate (BA 31) and uncorrected (p < .001) activation of the right medial frontal gyrus (BA 10) and bilateral middle temporal gyrus (BA 39) (See Figure 5 and Table 7).
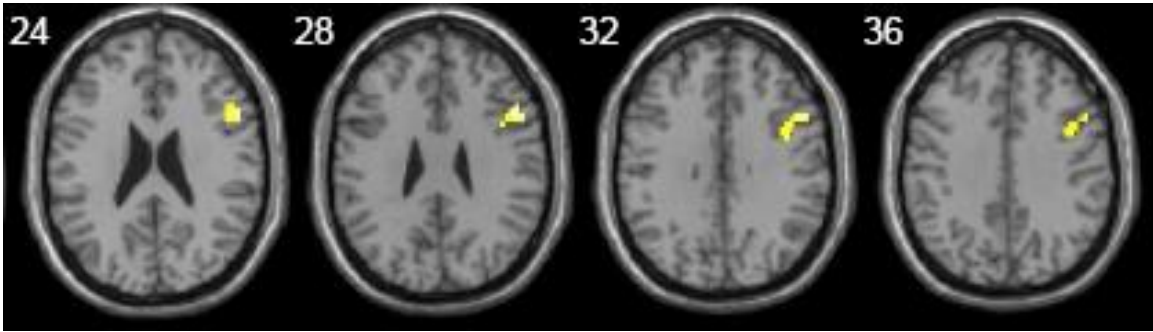
Figure 4: Activation for C2(+)>C1 contrast for combined study (MNI coordinates: x=50, y=15, z=as above).
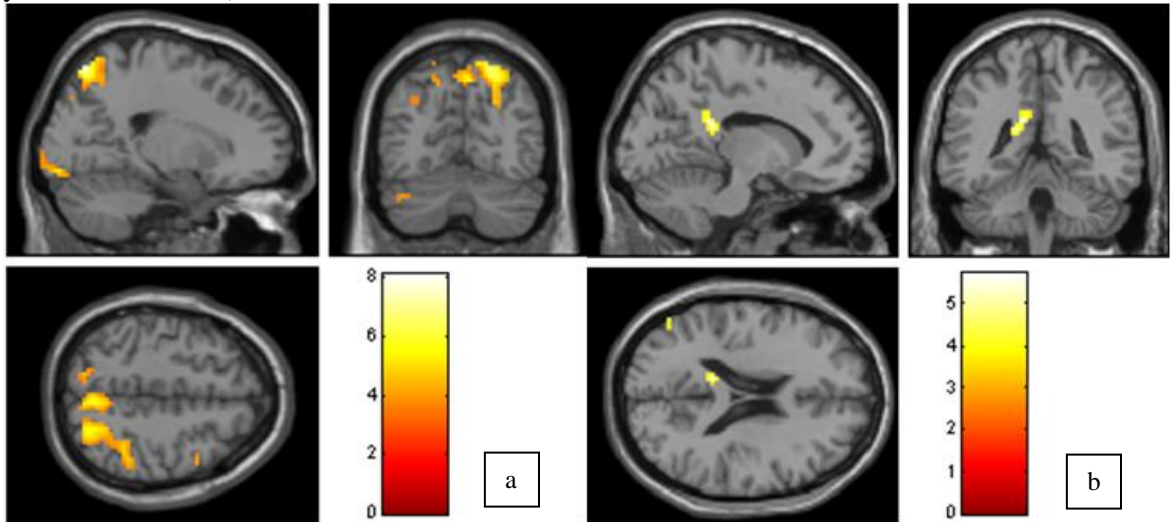

Figure 5: Comparison of activation for C2(+)>C1 contrast in: (a) self-paced design (MNI: 21, -69, 51) and (b) timed condition (MNI: -12, -42, 21).

Table 9: Results of the random-effect analysis in different contrasts

Regions of Activation	BA	Cluster voxel size (uncorrected value)	Talaraich coordinates			Z-score
Anatomical label			x	y	Z	
Color Relational-2+ > Color Relational-1 (Combined Study, N=42):						
Frontal						
R Middle frontal gyrus	9	71	50	16	27	4.47
R Precentral gyrus	9		42	7	33	3.57
Color Relational-2+ > Color Relational-1 (Study A, N = 22):						
Frontal						
R Inferior frontal gyrus	9	120	48	16	24	4.55*
R Middle frontal gyrus	6	48	30	20	57	4.11*
R Middle frontal gyrus	6		33	9	55	3.64*
R Middle frontal gyrus	6		45	8	49	3.63*
Parietal						
R Superior parietal lobule	7	446	21	-67	59	5.40*
R Precuneus	7		3	-62	50	4.50*
R Precuneus	7		24	-68	39	4.07*
L Precuneus	7	13	-15	-67	50	3.69*
L Angular gyrus	39	26	-30	-59	39	3.61*
Occipital						
R Middle occipital gyrus	18	177	30	-93	2	4.46*
R Middle occipital gyrus	18		21	-88	-6	4.12*
R Middle occipital gyrus	18		33	-93	13	3.95*
L Inferior occipital gyrus	18	167	-30	-90	-3	4.32*
L Cuneus	18		-24	-99	2.43	4.15*
L Inferior occipital gyrus	18		-42	-85	-3	4.14*
L Middle occipital gyrus	18	16	-9	-93	13	3.86*
Cerebellum						
L Cerebellum – Declive		21	-42	-71	-22	3.81*
Color Relational-2+ > Color Relational-1 (Study B, N = 20):						
Frontal						
R Medial frontal gyrus	10	19	3	53	3	3.71
Temporal						
L Cingulate gyrus	31	43	-12	-39	21	4.30
L Middle temporal gyrus	39	14	-53	-66	25	4.26
L Middle temporal gyrus	39	13	48	-63	28	4.03

All regions consist of at least 10 voxels with an uncorrected value of p < .001.
* significant at *p* < .05 False Discovery Rate correction

In the female subjects alone (*N* = 21, *M(age)* = 23.76), the 2(+)-relational trials revealed activation of the left precentral gyrus (BA 4) (uncorrected, p <.001) (See Figure 6a). 2(+)-relational trials for females in Study A showed additional activation (FDR

correction <.05) in bilateral middle occipital gyrus (BA 18/19), left cuneus (BA 18), left inferior occipital gyrus (BA 18), bilateral precuneus (BA 7), and right inferior parietal lobule (BA 40). Uncorrected results (p < .001) indicate additional activation of the left insula (BA 13). Study B did not reveal any unique activation for females at an uncorrected (p < .001) threshold (See Table 8).

Table 10: Results of the random-effect analysis in female subjects

Regions of Activation	BA	Cluster voxel size (uncorrected value)	Talaraich coordinates			Z-score
Anatomical label			x	y	z	
Color Relational-2+ > Color Relational-1 (Females only, Combined Study, N=21):						
Frontal						
L Precentral gyrus	4	10	-36	-23	62	3.28
Color Relational-2+ > Color Relational-1 (Females only, Study A, N =11):						
Parietal						
R Precuneus	7	209	24	-65	36	4.09*
R Inferior parietal lobule	40		36	-41	46	4.03*
L Precuneus	7	25	0	-65	50	3.68*
R Precuneus	7		3	-53	52	3.54*
Occipital						
R Middle occipital gyrus	18	157	33	-90	2	5.08*
R Middle occipital gyrus	18		27	-96	10	4.40*
R Middle occipital gyrus	19		39	-81	12	3.99*
L Cuneus	18	182	-27	-93	-1	4.57*
L Inferior occipital gyrus	18		-39	-87	-1	4.33*
L Middle occipital gyrus	18		-21	-96	5	4.15*
L Insula	13	10	-30	-40	24	3.48
Color Relational-2+ > Color Relational-1 (Females only, Study B, N =10):						
No uncorrected (p<.001) clusters above 10 voxels						

All regions consist of at least 10 voxels with an uncorrected value of p < .001.
* significant at *p* < .05 False Discovery Rate correction

Male subjects alone (*N* = 21, *M(age)* = 23.67) revealed 2(+)-relational complexity activation at the uncorrected *p*<.001 threshold in areas of the right inferior frontal gyrus (BA 9) (See Figure 6b). Study A contributes activations in the right inferior frontal gyrus (BA 9) and right superior parietal gyrus (BA 7) (Cluster-level correction < .05).

Additional activations were noted at the uncorrected (p < .001) threshold in the right middle frontal gyrus (BA 6), left decline of the cerebellum, and the left lingual gyrus (BA 18). Study B revealed cluster-level corrected (<.05) activation of the right angular gyrus (BA 39), and uncorrected (p<.001) activation in the culmen of the cerebellum (See Table 9).

Table 11: Results of the random-effect analysis for male subjects

Regions of Activation	BA	Cluster voxel size (uncorrected value)	Talaraich coordinates			Z-score
Anatomical label			x	y	Z	
Color Relational-2+ > Color Relational-1 (Males only, N=21):						
Frontal						
R Inferior frontal gyrus	9	34	42	7	30	3.92
Color Relational-2+ > Color Relational-1 (Males only, Study A, N =11):						
Frontal						
R Inferior frontal gyrus	9	33	42	7	27	4.77
R Inferior frontal gyrus	9		48	13	24	3.64
R Middle frontal gyrus	6	15	33	9	60	3.86
R Middle frontal gyrus	6		42	6	55	3.40
Parietal						
R Superior parietal lobule	7	30	24	-67	56	4.33
R Superior parietal lobule	7		27	-55	61	3.31
Temporal						
L Lingual gyrus	18	(24)	-6	-79	-6	3.22
Cerebellum						
L Cerebellum – declive		24	-24	-79	-14	3.57
Color Relational-2+ > Color Relational-1 (Males only, Study B, N =10):						
Parietal						
R Angular gyrus	39	35	53	-60	31	4.74
Cerebellum						
R Culmen		21	3	-59	-20	4.09
R Anterior lobe			9	-48	-25	3.79

All regions consist of at least 10 voxels with an uncorrected value of p < .001.
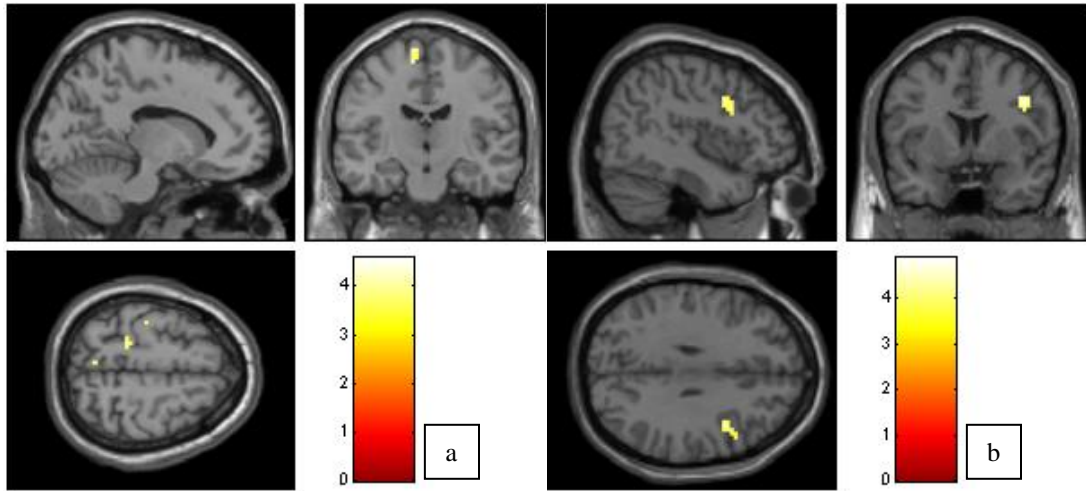
Figure 6: Relevant gender-specific activation for 2(+)-complexity trials: (a) Females (MNI coordinates: -12, -15, 60), (b) Males (MNI coordinates: 44, 6, 31).

Unique activation in males (uncorrected, p<.001) reveals regions of the right declive of the cerebellum, regions of the left cingulate (BA 32) and right precentral gyrus (BA 3/6) in Study A (See Figure 7a and Table 10). There were no significant activations for males in Study B.

Table 12: Results of the random-effect analysis contrasting gender

Regions of Activation	BA	Cluster voxel size (uncorrected value)	Talaraich coordinates			Z-score
Anatomical label			x	Y	z	
Relational Complexity: Males>Females, Combined Study						
Cerebellum						
R Declive		11	30	-56	-12	3.64
Relational Complexity: Males>Females, Study A:						
Frontal						
L Cingulate	32	24	0	30	26	4.21
R Precentral gyrus	6	42	45	-9	58	3.88
R Precentral gyrus	6		50	-6	53	3.81
R Precentral gyrus	3		50	-18	53	3.24

All regions consist of at least 10 voxels with an uncorrected value of p < .001.
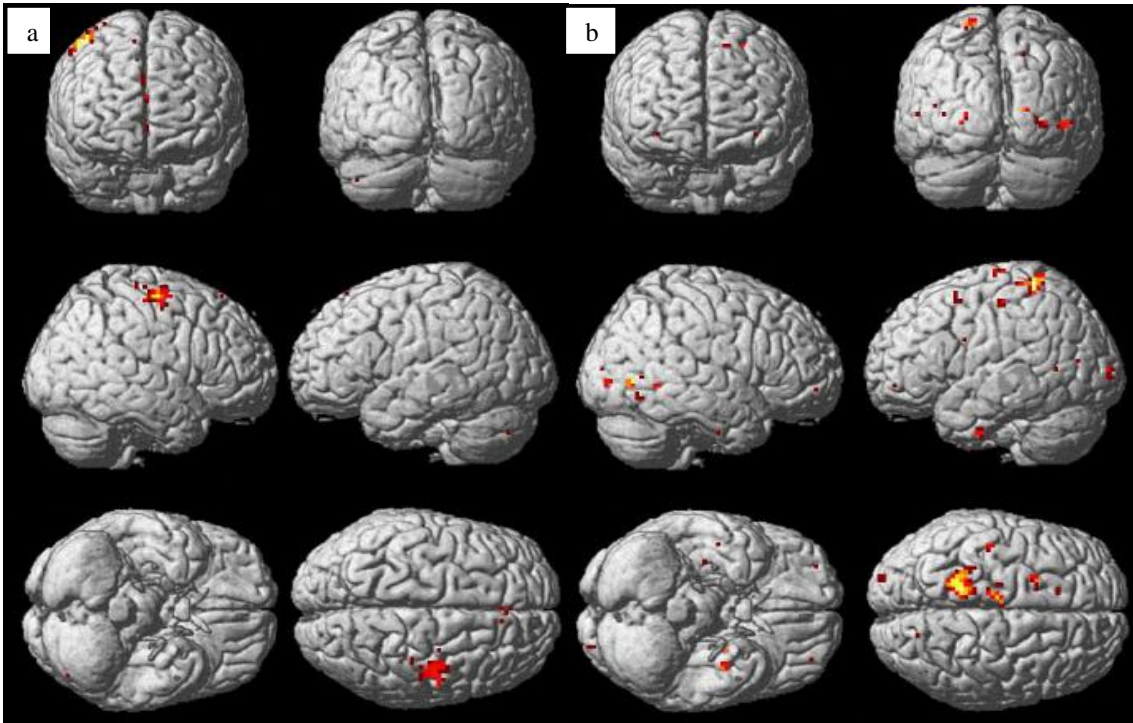
Figure 7: Results from Study A for (a) 2(+)-relational complexity in males, (b) 2(+)-relational complexity in females

Activation unique to female participants was more diffuse and robust. In particular, females uniquely activated regions (FDR corrected <.05) of the left superior, left middle, and left medial frontal gyrus (BA 6/32), left postcentral gyrus (BA 3), right cuneus (BA 7), right middle occipital gyrus (BA 18), left insula (BA 13), bilateral precuneus (BA 7), bilateral cingulate (BA 24/31/32), left thalamus, right precentral gyrus (BA 6), caudate tail, and bilateral middle temporal gyrus (BA 37/39). Female specific activation in Study A revealed regions of left caudate body, left postcentral gyrus (BA 3/5), right precentral gyrus (BA 6), and right caudate tail (cluster-level correction <.05). Additional activations were noted at the uncorrected (*p*<.001) threshold in left superior frontal gyrus (BA 6), right inferior temporal gyrus (BA 37), tail of the right caudate, bilateral anterior cingulate

(BA 24/32), right parahippocampal gyrus (BA 19), right cuneus (BA 17), right middle occipital gyrus (BA 19), left superior frontal gyrus (BA 6), right middle frontal gyrus (BA 8), and left middle temporal gyrus (BA 19) (See Figure 7b and Table 10). Female specific activation in Study B reveals regions of the bilateral caudate head, right medial frontal gyrus BA 6), left putamen, and left precuneus (BA 7).

Table 13: Results of the random-effect analysis contrasting gender

Regions of Activation	BA	Cluster voxel size (uncorrected value)	Talaraich coordinates			Z-score
Anatomical label			x	y	Z	
Relational Complexity – Females > Males, Combined Study:						
Frontal						
L Superior frontal gyrus	6	89	-24	8	52	4.76*
L Middle frontal gyrus	6		-33	5	49	3.55*
L Medial frontal gyrus	32		-15	14	46	3.53*
L Postcentral gyrus	3	94	-15	-38	65	4.54*
L Cingulate	31	(151)	-24	-45	35	3.77*
L Medial frontal gyrus	6	26	-9	-11	87	4.02*
L Thalamus		20	0	-23	15	3.96*
R Precentral gyrus	6	302	36	-4	33	3.93*
R Caudate – tail			24	-31	26	3.91*
L Cingulate gyrus	24	75	-9	-4	31	3.82*
L Cingulate gyrus	32		-18	13	32	3.59*
R Anterior cingulate	32	25	24	33	9	3.84*
Parietal						
L Precuneus	7	(151)	-9	-50	55	3.78*
R Precuneus	7	32	18	-53	47	3.86*
R Precuneus	7		18	-60	42	3.62*
Occipital						
R Cuneus	7	43	24	-75	7	4.15*
R Middle occipital gyrus	18		21	-84	12	3.35*
L Insula	13	151	-30	-37	21	4.02*
Temporal						
R Middle temporal gyrus	37	38	48	-52	0	3.77*
R Caudate – tail			36	-41	2	3.65*
L Middle temporal gyrus	39	17	-39	-72	15	3.66*
L Middle temporal gyrus	39		-48	-69	17	3.42*
Relational Complexity: Females>Males, Study A:						
Frontal						
L Superior frontal gyrus	6	19	-9	-14	-67	4.27
R Precentral gyrus	6	(168)	36	-4	33	3.76
R Anterior cingulate	32	13	24	36	12	3.62
L Superior frontal gyrus	6	10	-18	11	52	3.56
R Middle frontal gyrus	8	12	24	13	35	3.56
L Cingulate gyrus	24	32	-12	-4	31	3.52
L Cingulate gyrus	32		-18	7	33	3.52

	Parietal						
L	Insula	13	105	-27	-39	27	4.54
L	Caudate – body			-18	-28	26	3.82
L	Postcentral gyrus	3	94	-15	-38	63	4.19
L	Postcentral gyrus	5		-24	-41	60	4.15
L	Postcentral gyrus	3		-30	-32	62	3.40
R	Caudate – tail		168	24	-31	26	4.00
R	Caudate – tail		25	24	-37	10	3.66
	Occipital						
R	Cuneus	17	20	24	-75	7	3.57
R	Middle occipital gyrus	19		33	-81	4	3.34
	Temporal						
R	Inferior temporal gyrus	37	12	48	-67	1	3.87
R	Parahippocampal gyrus	19	31	42	-47	-3	3.62
R	Caudate – tail			36	-41	2	3.31
L	Middle temporal gyrus	19	17	-33	-61	11	3.43

Relational Complexity: Females>Males, Study B:

	Frontal						
R	Caudate – head		23	12	26	1	4.64
R	Medial frontal gyrus	6	18	3	11	44	3.82
L	Putamen		17	-21	20	-6	3.55
L	Caudate – head			-12	20	2	3.35
	Parietal						
L	Precuneus	7	11	-9	-50	55	3.47

All regions consist of at least 10 voxels with an uncorrected value of p < .001.
* significant at *p* < .05 False Discovery Rate correction

DISCUSSION

The initial goal of this study was to examine the gender-specific systems present during relational complexity by combining results from two independent fMRI tasks to exploit a larger sample size. This relationship was however further defined post-hoc by examining the influence of task parameters on trial performance and brain activation. For that reason our results will refer only to the former gender discrepancy hypotheses with the task discrepant results being discussed secondarily.

The present analysis aimed to identify the neural underpinnings of the reasoning process in the context of color relational complexity for males and females. As predicted, subjects, regardless of gender, solved the 1-relational complexity trials significantly faster and more accurately than the 2(+)-relational complexity trials. This demonstrates that the introduction of relational complexity induces a consistent behavioral response when faced with increased difficulty and cognitive demand. Examination of behavioral results considering additional groupings also reveals greater accuracy and faster reaction times during 1-complexity trials regardless of gender or task design. These finding are consistent with the relational complexity model and support the validity of our color relational complexity (CRC) task design. Additionally, this suggests that the proposed complexity introduced by the 2(+)-relational complexity trials is maintained in both study conditions and is observed in both genders.

We first characterized the activation specific to color relational complexity at the 2(+)-complexity level to include all subjects regardless of gender or study. As hypothesized, a large cluster of activation was noted in the right middle frontal gyrus (a region consistent with the DLPFC), likely supporting spatial processing, fluid reasoning, and working memory (Chirstoff, et al., 2001; Kroger, et al., 2002). Contrary to our hypothesis, however, there was an absence of left anterior PFC activity in the combined group activation map for 2(+)-complexity. This absence of left lateralized activity may result from the limited parametric variation we introduced in our design. According to Kroger and colleagues (2002), the left anterior PFC activation was produced in response to their highest complexity trials, which were categorized from 3-4 levels or relational complexity. Our relational complexity design, on the other hand sought to only characterize the differentiation between only two levels of complexity (1 and 2), which may not provide sufficient cognitive demand to produce hemispheric shifts.

It should also be considered that the unsolicited ambiguity present in our color-relational complexity stimuli resulted in an activation pattern not able to be attributed to a single level of complexity. Our representation of color relational complexity stimuli, although prevalent in the literature and historically validated, may not actually represent an accurate portrayal of a 1-relational and 2-relational complexity task. Although we modeled our self-made CRC stimuli on previous research, based on a post-hoc inter-rater, item-by-item analysis we have determined that those trials designated as 2-relational may in fact be better qualified as 2+, as some have been interpreted to rely on 3 aspects of integration for accurate solutions. This ambiguity in the trials deemed as

having a 'higher complexity' (2+) therefore requires an analysis that considers 2-complexity and 3-complexity trials as one group and does not allow for the differentiation present in the study conducted by Kroger and colleagues (2002) which produced the left lateralized PFC activity we had hypothesized.

Additionally the graphical components of the RC design may be inadvertently presenting tasks that are interpreted as 12-item prompts, when it is typically assumed to be an 8-item prompt. This may result in misinterpretation of the correct number of complexity levels and therefore not be a true examination of 1-relational versus 2-relational. However, we are confident that we have characterized a distinction between a lower level of relational complexity, as represented by 1-relational trials and a higher level of relational complexity, as represented by 2(+)-relational trials. Evidence for this distinction, beyond historical precedent, is supported by the behavioral performance of our task that indicates significantly reduced accuracy and increased reaction time for those trials designated as being more complex. Furthermore, since it has been determined that humans have the ability to integrate up to four dimensions of variation without the need for external memory aids (Halford, Wilson, & Phillips, 1998), we know that, given the high degree of accuracy in our behavioral model, we presented a task within the capabilities of our participants. Therefore, even if the number of true relations present in our trials is questionable, the underlying concept we attempted to measure (i.e., relational complexity) is maintained.

In an effort to better characterize the influence of each task design on the phenomenon of relational complexity, we rendered the relevant activation specific to each study. Doing so revealed a large divergence of activity and localization. Notably, in the self-paced condition (Study A), subjects relied heavily on regions of the superior parietal and precuneus, middle and inferior occipital gyri, and middle and inferior frontal gyri. This superior parietal activation is consistent with previous literature employing the use of non-verbal reasoning tasks (Hugdahl, Thomsen, & Ersland, 2006; Clements-Stevens, Rimrodt & Cutting, 2009; Jordan, et al. 2002) suggesting its role in visuospatial processing, with the precuneus specifically implicated in motor imagery (Hanakawa, et al., 2003), mental rotation (Suchan, et al., 2002), task load and spatial processing (Kalbfleisch, Van Meter, & Zeffiro, 2007) and self-consciousness and self-centered mental imagery strategies (Cavana & Trimble, 2006). The activation noted in the middle occipital region has been implicated in visual attention (Martinez, et al., 1999) whereas middle and inferior frontal gyri have been shown to play a role in relational complexity and planning (Christoff, et al., 2001; Kalbfleisch, et al., 2007) and support cognition during deductive reasoning (Osherson, et al. 1998).

Activation observed in the timed condition (Study B) yielded less robust results to include areas of the cingulate and medial frontal lobe. As above, this frontal activity is consistent with prior research, whereas the cingulate is most often implicated in emotion and conflict resolution (Botvinick, Cohen, & Carter, 2004).

The variation noted in the activation maps between studies may also be dependent on the constraints placed on the task designs themselves, rather than their content. In Study

A the subjects were afforded an open-ended, self-paced design whereby they could solve the CRC trial in relatively unpressured conditions. In Study B, subjects were placed under finite timing constraints. Study B required a very swift solution and subsequent response, where visual and audible cues were used to alert the subject to respond. This time period was determined from behavioral data in Study A, to indicate the appropriate amount of maximum time that would be given for subjects to solve the CRC problems in Study B. It is possible that cognition under these two paradigms looks very different because of the overshadow provided by the cued response employed in Study B. This suggestion is supported based on the performance levels between the studies, whereby Study A performance, which afforded much more time for solving the trial, elicited greater accuracy across gender and complexity. Despite this apparent variability in task design, the performance trends between gender and complexity remained. In other words, the significant behavioral findings noted in Study A were also revealed in Study B.

It is possible that the greater activation noted in Study A as compared to Study B can be attributed to the difference in our sample size (22 subjects vs. 20 subjects, respectively). This is, however, unlikely, as we have characterized robust activation in sample sizes with less than 20 subjects. Based on participant's behavioral performance there is, as expected, a higher level of accuracy in the self-paced condition, which suggests that the differences noted above were performance-based and not condition-based. In support of an assumption that this pattern is condition-based, we have attempted to control for large variance in the performance between studies by thresholding the inclusion criteria to characterize only those individuals that maintained a high degree of

accuracy in all conditions (>85%), as well as including only those trials that were solved accurately in our imaging analysis. This had the effect of producing a varying number of useable trials for each individual and subsequent group analysis. Given future opportunity, it would be important to covary for performance within the fMRI analysis, as well as, maintain equal numbers of trials between our test conditions. Since the condition-based analysis was secondary to the gender-based analysis we did not match trials with the intention of comparing Study A versus Study B, but rather for comparing Male versus Female. Therefore as it pertains to the condition-based results we must concede that the variance in our statistical power, dependent on number of trials per participant, is a limitation to our results. There is an additional possibility that the amount of time subjects were attuned to the task in each condition were inherently varied, therefore allowing participants the ability to concentrate and reason through the self-paced design while not being afforded the same luxury on the timed condition. This was controlled for in the analysis, by relying on a single-event to define our response (moment of button pressing) rather than relying on an epoch-related response thereby evening out the contribution that additional processing might incur. From a performance perspective, this reveals the potential influence that pressured/timed conditions may place on underlying neural systems.

Despite these caveats, we propose that the introduction of time constraints on spatial reasoning, in this case during color relational complexity, results in a more parsimonious, albeit self-conscious process, whereas the self-paced design elicits activation typically

noted in examination of less time-dependent spatial tasks in the superior parietal and middle/inferior occipital regions (Clements-Stevens, et al., 2009).

Since the influence of gender on the performance and activation noted during relational complexity was the primary motivation for the study, several contrasts were conducted to gain a more complete understanding of supporting neural systems. Both male and female activation maps for relational complexity were examined independently. We contrasted the 2(+)-complexity trials with the 1-complexity trials for all 21 females and then contrasted the 2(+)-complexity trials with 1-complexity trials for all 21 males. This process garnered two independent gender-defined group maps representing processing at the 2(+)-complexity level, which were further interrogated to examine influence of task design within these independent gender maps. Gender discrepancy was also explored directly by subtracting the 2(+)-complexity female maps from the 2(+)-complexity male maps, and vice versa, both of which were also separated by task design.

Gender discrepancy was revealed in the activation maps we considered, as was the influence of task design. Consistent with our hypothesis, male-specific activation was more pronounced in the posterior/parietal regions than in female participants. Activation of the spatial processing-associated regions of the superior parietal lobule and lingual gyrus in the male brain maps were notably absent in the female brain maps. Complementary to this hypothesis, much of the male activation derived from cerebellar activity to include bilateral declive, right anterior lobe, and right culmen, a finding that is not found in female subjects. This male-specific activation may suggest reliance on the cerebellum based on its association with higher-order functioning (Schmahmann &

Sherman, 1998; Kalbfleisch, et al., 2007), as well as, indications that left cerebellar excision results in visuo-spatial deficits (Wallesch & Horn, 1990). Furthermore, the consistent activation of the cerebellum in the 2-relational trails may reflect compensatory recruitment in the face of uncertainty (Blackwood, et al., 2004). Therefore, the increased activation of the cerebellum in the context of relational complexity may suggest a need for visuo-spatial organization and processing to guide the timing and approximation of unconfirmed complex decisions and eventual responses, as evidenced by the increased accuracy and slower reaction times noted in our sample.

Our hypotheses pertaining to male reliance on bilateral activation and female reliance on right hemispheric activation were not supported and our results were in opposition to this. Generally, regardless of the task condition, females tended to recruit bilateral regions, such as the caudate, cingulate, middle temporal, and precuneus, whereas males tended to rely primarily on regions of the right hemisphere.

While some of our findings were consistent with proposed hypotheses, there were several other trends observed in our results that were unexpected. The gender comparison for the CRC task revealed a peculiar interaction. Male participants were significantly more accurate than the female participants; however females were significantly faster to provide an accurate response. Although this performance discrepancy was not represented by a large effect size, the functional contrasts indicated that women uniquely activated a greater number of cortical regions during relational complexity reasoning than men. This likely suggests that men rely on a more parsimonious approach to the relational complexity demands at the risk of decreased efficiency. Conversely, this

suggests that the neural networks employed by women likely resulted in more time-efficient processing at the cost of decreased accuracy. Specifically women exhibited greater activation in frontal lobe regions of the superior and middle gyri, as well as subcortical areas of the cingulate, thalamus, and caudate. The network of frontal lobe activity may represent simultaneous information processing, leading to a rapid reaction time, while the limbic associated regions may result in increased interference, attributing to the noted decrease in performance. The caudate may also serve to mediate spatial working memory for planning and execution of motor action (Postle & D'Epoosito, 1999).

Generally, females elicited greater activation from neural systems affiliated with affective processing, suggesting extraneous processing providing support while processing task-related demands. A suggestion proposed by Birenbaum, Kelly, and Levi-Keren (1994) is that non-verbal reasoning deficits in women (i.e., 2D mental rotation) may be the result of their association with personality traits, such as "obsessive perfectionism" or "cautiousness", suggesting self-conscious perceptual influences. Activations noted above may support this theory, but there is another potential influence we propose that is not mutually exclusive.

Based on the decreased accuracy of our female participants, coupled with the female-specific activation of emotion-regulatory brain regions, we suggest that this gender discrepancy during our relational complexity task may stem from a phenomenon known as stereotype threat (ST). Stereotype threat refers to the added pressure placed on an individual when their resultant performance may confirm a negative stereotype about

their ingroup (i.e., females), thereby negatively affecting their performance (Schmader & Johns, 2003). These threats to performance are typically exhibited in female participants during typical gender-discrepant tasks (i.e., math, spatial reasoning) and/or amidst non-group members (Sekaquaptewa & Thompson, 2003; Inzlitch & Ben-Zeev, 2000; Inzlitch & Ben-Zeev, 2003; McGlone & Aronson, 2006).

Schmader and Johns (2003) suggest that performance discrepancies between groups be better addressed through situational explanations, and not through group ability. Such an explanation relying on situational context is supported by the gender discrepancy present in our results. Female's are required to undergo a more involved and invasive pre-scan routine prior to their imaging session. This procedure, including a mandatory pregnancy test and instructions to not wear typical undergarments (i.e., wire bra) or any make-up, is unique to the female participants only, and therefore provides an environmental context and potential pre-emptive stressor that is not characteristic of the male participants. Although a post-interview to determine anxiety levels, performance expectations, and/or performance evaluation was not conducted, these influences cannot be ruled out. Despite our inability to identify these traits in our female participants, impaired performance on a task does not require conscious awareness of performance expectancies, indicating that the ST effects may be the result of implicit processes (Sekaquaptewa & Thompson, 2003; Schmader & Johns, 2003; Major & O'Brien, 2005),. We submit that this phenomenon is characterized in our functional neuroimaging results, implicating the involvement of frontal and limbic regions during reasoning in our female participants. This limbic activation, not seen in men, suggests that the female participants

are likely overtaxing their working memory capacity with extraneous anxiety suppressing resources (Schmader & Johns, 2003) resulting in a significant reduction in CRC accuracy.

Compared with the results outlined in our current study, Krendl, Richeson, Kelley, and Heatherton (2008) characterized a similar pattern in female participants undergoing fMRI during a math task. The typical problem solving regions of the PFC were activated as expected, but when ST was induced, women also activated regions of the cingulate cortex, thought to recruit social and emotional processing thereby usurping the successful recruitment of other neural networks which have aided in accurate problem solving. Schmader, Johns, and Forbes (2008), suggest that this ST disrupts performance by PFC impairment due to a physiological stress response, active performance monitoring and self-regulatory, negative-emotion suppression. Therefore as applied to our results, we see that females were more likely to recruit areas of the frontal lobe to include the superior and middle gyri, as well as subcortical areas of the cingulate, thalamus, and caudate, the latter of which seizes central task processing resources for anxiety suppression contributing to suboptimal performance (Johns, Inzlicht, & Schmader, 2008).

Also note that although the mean age for our participants in Study A vs. Study B (*M=24.2* and *M*=22.2, respectively) and Male vs. Female (*M*=23.7 and M=23.8, respectively) were nearly identical, four of our individuals were older than 30 years of age. Based on a literature review we found no conclusive evidence for major neural changes of note within this developmental span as research of this type focuses on the young (18-30 years old) vs. the old (60 and older) (Persson, Lustig, Nelson, & Reuter-

Lorenz, 2007; Reuter-Lorenz, Jonides, Smith, Hartly, Miller, Marshuetz, & Koeppe, 2000; Nyberg, Sandblom, Jones, Neely, Petersson, Ingvar, & Backman, 2003; Rypma, Prabhakaran, Desmond, & Gabrielli, 2001). Although this range was not indicative of our participant ages, we hoped to solidify our findings by accounting for the variability any age discrepancy may contribute. This was addressed secondarily through an alternate functional image analysis covaried by age. Results of the covaried activation within each contrast revealed similar activations to our original results, albeit with reduced power, for several of our contrasts. For comparative reference we would like to highlight several relevant brain regions present in both analyses, however the results of this ancillary covaried analysis will not be discussed in equally exhaustive detail as our original analysis. For example, the phenomenon of relational complexity within the self-paced condition continues to implicate regions associated with visuospatial processing and reasoning to include the middle frontal lobe, the superior parietal lobule and precuneus, the inferior occipital lobe, and the cerebellum. Another notably consistent finding, following our covaried analysis, reveals significant activation of the regulatory regions of the limbic area to include the thalamus and cingulate, as well as, a region of the middle frontal gyrus.

Based on the internal consistency of our behavioral data and the overlap of our activation maps with previous literature, we determine that we have accurately captured the phenomenon of relational complexity. The influence of gender on this higher order reasoning task has been in debate and here we have identified additional evidence of divergent neural processing. Herein we determined that while women were significantly

faster to accurately solve the trials, this efficiency came at the cost of decreased performance. Despite the slight effect sizes noted, this difference accounted for the implication of different neural systems across genders. Men were shown to rely on a more conserved parsimonious neural system for relational complexity demands. Women were recruiting prefrontal and limbic regions thereby enhancing contextual awareness at the cost of an increase in working memory. This discrepancy was further elucidated via examination of the task within their varied designs. The self-paced condition elicited robust patterns of activation consistent with typical non-verbal reasoning and visual-spatial processing, whereas the timed condition elicited highly conserved activation consistent with self-regulation monitored through the cingulate. Based on these results future neuroimaging research would be best served to examine not only the gender component, but also the potential external pressures posed on the female participants not imposed on the male participants (i.e., removal of make-up, dressing down for the scan day, mandatory pregnancy test). It may also be important to conduct post-scan interviews to better chronicle the potential feelings that participants were experiencing during scanning (i.e., "I felt judged", "I really wanted to do a good job"). Additionally note that these internal or external discrepancies imposed by gender might be further exacerbated or confounded given a particularly demanding task or pressured situation. Additional studies would need to be conducted in order to determine whether the results found here could be practically applied to the educational field in order to predict and optimize the best performance from their individuals based on the potential gender influence and task paradigm.

In summary, males tended to show activity more consistent with visuospatial networks (e.g., cerebellum, right superior parietal, middle frontal) whereas females tended to rely on spatial attention, emotional regulation and conflict resolution, and spatial working memory regions (e. g., superior frontal, cingulate, middle occipital, caudate). Consistent with Clements and Stevens (2009) these findings may suggest a male reliance on visual imagery underlying the superior performance on color relational complexity tasks with females relying on the convergent function of several brain regions to support their increased efficiency, at the risk of competing regulatory regions compromising accuracy. The contextual findings herein, supported by both behavioral and neural discrepancies between the genders, was unexpected. Nevertheless, although we had initially sought to characterize the gender-unique procedural processes that could contribute to divergent performance, we instead uncovered an environmental influence, which requires further consideration, especially in the context of neuroimaging pre-scan procedures and subsequent group analysis.

REFERENCES

Bell, E. C., Willson, M. C., Wilman, A. H., Dave, S., & Silverstone, P. H. (2006). Males
and females differ in brain activation during cognitive tasks. *Neuroimage, 30,*
529-538. doi: 10.1016/j.neuroimage.2005.09.049

Birenbaum, M., Kelly, A. E., & Levi-Keren, M. (1994). Stimulus features and sex
differences in mental rotation test performance. *Intelligence, 19*, 51-64. doi:
10.1016/0160-2896(94)90053-1

Blackwood, N., Ffytche, D., Simmons, A., Bentall, R., Murray, R., & Howard, R. (2004).
The cerebellum and decision making under uncertainty. *Cognitive Brain
Research, 20*, 46-53. doi: 10.1016/j.cogbrainres.2003.12.009

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior
cingulate cortex: An update. *Trends in Cognitive Sciences, 8,* 539-546.
doi:10.1016/j.tics.2004.10.003

Butler, T., Imperato-McGinley, J., Pan, H., Voyer, D., Cordero, J., Zhu, Y-S., Stern, E.,
& Silbersweig, D. (2006). Sex differences in mental rotation: Top-down versus
bottom-up processing. *Neuroimage, 32,* 445-456. doi:
10.1016/j.neuroimage.2006.03.030

Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: A review of its functional

    anatomy and behavioural correlates. *Brain, 129,* 564-583. doi:

    10.1093/brain/awl004

Christova, P. S., Lewis, S. M., Tagaris, G. A., Ugerbil, K., & Georgopoulos, A. P. (2008).

    A voxel-by-voxel parametric fMRI study of motor mental rotation: hemispheric

    specialization and gender differences in neural processing efficiency.

    Experimental Brain Research, 189, 79-90. doi: 10.1007/s00221-008-1405-x

Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, S. H., Mostofsky, S. H., Pekar, J.

    J., Denckla, M. B., & Cutting, L. E. (2006). Sex differences in cerebral laterality

    of language and visuospatial processing. *Brain and Language, 98,* 150-158. doi:

    10.1016/j.bandl.2006.04.007

Clements-Stephens, A. M., Rimrodt, S. L., & Cutting, L. E. (2009). Developmental sex

    differences in basic visuospatial processing: Differences in strategy use?

    Neuroscience Letters, 449, 155-160. doi: 10.1016/j.neulet.2008.10.094

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., &

    Gabrieli, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational

    integration during reasoning. *Neuroimage, 14*, 1136-1149. doi:

    10.1006/nimg.2001.0922

Deeb, S. S., (2004). Molecular genetics of color-vision deficiencies. *Vision Neuroscience*,

    21, 191-6. doi: 10.1017/S0952523804213244

Donaldson, D. I., & Buckner, R. L. (2001). Effective paradigm design. In P. Jezzard, P.

    M. Matthews, & S. M. Smith (Eds.), *Functional MRI: An Introduction to Methods*

    (pp. 178-179). New York: Oxford University Press.

Fennema, E. & Sherman, J. (1977). Sex-related differences in mathematics achievement,

    spatial visualization and affective factors. *American Educational Research*

    Journal, 14, 51-71. doi: 10.3102/00028312014001051

Freire, L., & Mangin, J. F. (2001). Motion correction algorithms may create spurious

    brain activations in the absence of subject motion. *Neuroimage, 14*, 709-722. doi:

    10.1006/nimg.2001.0869

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J-P., Frith, C. D., & Frackowiak, R.

    S. J. (1994). Statistical parametric maps in functional imaging: A general linear

    approach. *Human Brain Mapping, 2,* 189-210. doi: 10.1002/hbm.460020402

Gill, H. S., & O'Boyle, M. W. (1997). Sex differences in matching circles and arcs:

    Preliminary EEG investigation. *Laterality, 2,* 33-48. doi: 10.1080/713754253

Gur, R. C., Alsop, D., Glahn, D., Petty, R., Swanson, C. L., Maldjian, J. A., Turetsky, B.

    I., Detre, J. A., Gee, J. & Gur, R. E. (2000). An fMRI study of sex differences in

    regional activation to a verbal and a spatial task. *Brain and Language*, *7*, 157-170.

    doi: 10.1006/brln.2000.2325

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by

    relational complexity: Implications for comparative, developmental, and cognitive

    psychology. *Behavioral and Brain Sciences, 21,* 803-865. doi:

    10.1017/S0140525X98001769

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Gelderen, P. V., & Hallett, M.

    (2003). Functional properties of brain areas associated with motor execution and

    imagery. *Journal of Neurophysiology, 89,* 989-1002. doi: 10.1152/jn.00132.2002

Hugdahl, K., Thomsen, T., & Ersland, L. (2006). Sex differences in visuo-spatial

    processing: An fMRI study on mental rotation. *Neuropsychologia, 44,* 1575-1583.

    doi: 10.1016/j.neuropsychologia.2006.01.026

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why

    females are susceptible to experiencing problem-solving deficits in the presence

    of males. *Psychological Science, 11,* 365-371. doi: 10.1111/1467-9280.00272

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in

    private? The implications of threatening environments on intellectual processing.

    Journal of Educational Psychology, 95, 796-805. doi: 10.1037/0022-

    0663.95.4.796

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource

    depletion: Examining the influence of emotion regulation. *Journal of Experimental*

    Psychology: General, 137, 691-705. doi: 0.1037/a0013834

Jordan, K., Wustenburg, T., Heinze, H-J., Peters, M., & Jancke, L. (2002). Women and

    men exhibit different cortical activation patterns during mental rotation.

    Neuropsychologia, 40, 2397-2408. doi: 10.1016/S0028-3932(02)00076-3

Kalbfleisch, M. L. (2004). Functional neuroanatomy of talent. *The Anatomical Record,*

    227B, 21-36. doi: 10.1002/ar.b.20010

Kalbfleisch, M.L., DeBettencourt, M.T., Roberts, J.L., Banasiak, M.B., Kopperman, R., Halavi, M. (in submission). Environmental influence on neural systems of relational complexity.

Kalbfleisch, M.L., VanMeter, J.W., Zeffiro, T.A. (2007). The influences of task difficulty and response correctness on neural systems supporting fluid reasoning. *Cognitive Neurodynamics, 1,* 71-84. doi: 10.1007/s11571-006-9007-4

Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat: A functional magnetic imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science, 19,* 168-175. doi: 10.1111/j.1467-9280.2008.02063.x

Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex, 12*, 477-485. doi: 10.1093/cercor/12.5.477

Lancaster, J.L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A.C., Toga, A. W., & Mazziotta, J. C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping, 5,* 238-242. doi: 10.1002/(SICI)1097-0193(1997)5:4<238::AID-HBM6>3.0.CO;2-4

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach Atlas labels for functional brain mapping. *Human Brain Mapping, 10,* 120-131. doi: 10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8

Major, B., & O'Brien, L. T. (2005). The social psychology of stigma. *Annual Review of Psychology, 56,* 393-421. doi: 10.1146/annurev.psych.56.091103.070137

Martinez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., Wong, E. C., Hinrichs, H., Heinze, H. J., & Hillyard, S. A. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience, 2,* 364-369. doi: 10.1038/7274

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27,* 486-493. doi: 10.1016/j.appdev.2006.06.003

Naglieri, J. A. (1997). Naglieri nonverbal ability test. San Antonio, TX: The Psychological Corporation.

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., & Backman, L. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *Proceedings of the National Academy of Sciences, 100,* 13728-13733. doi: 10.1073/pnas.1735487100

Osherson, D., Perani, D., Cappa, S., Schnur, T., Grassi, F., & Fazio, F. (1998). Distinct

    brain loci in deductive versus probabilistic reasoning. *Neuropsychologia, 36,* 369-

    376. doi: 10.1016/S0028-3932(97)00099-7

Penny, W. D., Holmes, A. P., & Friston, K. J. (2003). Hierarchical Modeling. In

    Frackowiak,R. S. J., Friston, K. J., Frith, C., Dolan, R., Price, C. J., Zeki, S.,

    Ashburner, J., & Penny, W. D. (Eds.) *Human Brain Function* (pp. 1-20).

    Retrieved from http://www.fil.ion.ucl.ac.uk/spm/doc/biblio/Author/HOLMES-

    AP.html

Persson, J., Lustig, C., Nelson, J. K., & Reuter-Lorenz, P. A. (2007). Age differences in

    deactivation: A link to cognitive control? *Journal of Cognitive Neuroscience, 19,*

    1021-1032. doi: 10.1162/jocn.2007.19.6.1021

Prabhakran V., Rypma, B., & Gabrieli J. D. (2001). Neural substrates of mathematical

    reasoning: A functional magnetic resonance imaging study of neocortical

    activation during performance of the necessary arithmetic operations test.

    Neuropsychology, 15, 115-127. doi: 10.1037//0894-4105.15.1.115

Raven J. C. (1947). Raven's progressive matrices. San Antonio, TX: Psychological

    Corporation.

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartly, A., Miller, A., Marshuetz, C., &

    Koeppe, R. A. (2000). Age differences in the frontal lateralization of verbal and

    spatial working memory revealed by PET. *Journal of Cognitive Neuroscience, 12,*

    174-187. doi: 10.1162/089892900561814

Rypma, B., Prabhakaran, V., Desmond, J. E., & Gabrieli, J. D. E. (2001). Age differences in prefrontal cortical activity in working memory. *Psychology and Aging, 16,* 371-384. doi: 10.1037//0882-7974.16.3.371

Schmader, T. & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440-452. doi: 10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115,* 336-356. doi: 10.1037/0033-295X.115.2.336

Schmahmann, J. D., & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain, 121*, 561-579. doi: 10.1093/brain/121.4.561

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology, 39,* 68-74. doi: 10.1016/S0022-1031(02)00508-5

Suchan, B., Yaguez, L., Wunderlich, G., Canavan, A. G. M., Herzog, H., Tellmann, L., Homberg, V., & Seitz, R. J. (2002). Hemispheric dissociation of visual-pattern processing and visual rotation. *Behavioral Brain Research, 136,* 533-544. doi: 10.1016/S0166-4328(02)00204-8

Thomsen, T., Hugdahl, K., Ersland, L., Barndon, R., Lundervold, A., Srnievoll, A. I., Roscher, B. E., & Sunberg, H. (2000). Functional magnetic resonance imaging (fMRI) study of sex differences in a mental rotation task. *Medical Science Monitor, 6,* 1186-1196. doi: 10.1016/S0304-3940(03)00406-3

Unterrainer, J. M., Ruff, C. C., Rahm, B., Spreer, J., Schwarzwalkd, R., & Halsband, U. (2005). The influence of sex differences and individual task performance on brain activation during planning. *Neuroimage, 24*, 586-590. doi: 10.1016/j.neuroimage.2004.09.020

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250-270. doi: 10.1037/0033-2909.117.2.250

Wallesch, C-W., & Horn, A. (1990). Long-term effects of cerebellar pathology on cognitive functions. *Brain and Cognition, 14*, 19-25. doi: 10.1016/0278-2626(90)90057-U

Waltz, J. A., Knowlton, B. A., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santoa, M., Thomas, C. R., & Miller, B. L. (2002). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10,* 119-125. doi: 10.1111/1467-9280.00118

Weiss, E., Siedentopf, C. M., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C., Golaszewski, S., Felber, S., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in brain activation pattern during a visuospatial cognitive task: A functional magnetic resonance imaging study in healthy volunteers. *Neuroscience Letters, 344*, 169-172. doi: 10.1016/S0304-3940(03)00406-3

CURRICULUM VITAE

Joshua M. Roberts graduated from Maple Shade High School, Maple Shade, NJ, in 1997. He received his Bachelor of Arts in Psychology from the Richard Stockton College of NJ in 2001. He has been employed in Washington, DC as a supervisory manager for adults with developmental disabilities for five years and expects to receive his Ph.D. in Biopsychology from George Mason University in 2012.