

Sex Differences in Intelligence Areas and Response Time Tasks

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation with honors research distinction in Psychology in the undergraduate colleges of The Ohio State University

by

Natalie Bennett

The Ohio State University

June 2011

Project Advisor: Assistant Professor Fabio Leite, Department of Psychology

### **Acknowledgments**

We thank Lauren McKinley for data collection on the Lima campus of the Ohio State University.

**Abstract**

Previous research has shown that although there may not be sex differences in general intelligence, males tend to outperform females in the areas of spatial ability and mechanical reasoning, whereas females have an advantage in verbal ability and perceptual speed. Following these findings of sex differences in particular areas of intelligence, the present study tested for sex differences in participants' performance in response time tasks of different levels of complexity. We collected data from thirty-five Ohio State University students. In the first session, participants were administered the Wechsler Adult Intelligence Scale IV (WAIS-IV) to produce IQ estimates. In accordance with previous research, we predicted that females would have higher scores on verbal ability and perceptual speed WAIS-IV subtests. In the second part of the study, participants completed four response time (RT) tasks of different levels of complexity. We hypothesized that males would outperform females on the signal detection task (the simplest of the RT tasks), whereas females would outperform males on the lexical decision task (reliant on verbal ability). Results showed that males did indeed have faster RTs than females in the signal detection task. Additionally, male and female RTs on this signal detection task correlated differently with subtests of the WAIS-IV. However, no sex differences were found on any of the other three RT tasks. Furthermore, no sex differences were found on any subtests of the WAIS-IV. These findings show that although sex differences may not be obvious in intelligence or RT performance, males and females may use different resources to complete the same task.

### Sex Differences in Intelligence Areas and Response Time Tasks

It has long been argued that there are sex differences in intelligence, that perhaps men possess greater intellectual ability than women. The feminist movement in the late twentieth century caused the general population to question if such differences exist. The study of sex differences in intelligence is meaningful to psychologists because the existence of such differences can have applications in various aspects of the people's daily lives, such as the educational system (Gras, Bordoy, Ballesta & Berna, 2010), athletics (Spierer, Petersen, Duffy, Corcoran & Rawls-Martin, 2010), and driving a car (Silverman, 2006). This thesis explores sex differences in certain measures of intelligence and in some cognitive tasks driven by sensory evidence. In what follows, intelligence is operationally defined and previous findings about sex differences in intelligence, perceptual speed, and response time are discussed.

In the present study, general intelligence was assumed to be represented by a single number, the intelligence quotient (IQ) as measured by the Wechsler Adult Intelligence Scale IV (WAIS-IV; Wechsler, 2008). Such a general intelligence factor, often referred to simply as *g*, has been shown to be an underlying component of mental ability (Johnson, Bouchard Jr., Krueger, McGue & Gottesman, 2004). Johnson et al. (2004) found that three independent compilations of cognitive tests (i.e., the Comprehensive Ability Battery, the Hawaii Battery with Raven's Progressive Matrices, and the Wechsler Adult Intelligence Scale) had near perfect correlations with each other, demonstrating that they all measured an overall general intelligence of their participants. Looking for sex differences in *g*, some researchers have found a slight male advantage in overall IQ (Nyborg, 2005; Lynn, 1998). However, most agree that there is no sex difference in general intelligence (e.g., Kaufman, McLean & Reynolds, 1988; Camarata &

Woodcock, 2006; Halpern & LaMay, 2000), and that, in fact, IQ should not show any significant difference across gender (Wechsler, 1981).

Although sex differences may not exist in *g*, males and females perform differently in specific areas of intelligence. For example, males usually outperform females in the areas of spatial ability and mechanical reasoning; females, on the other hand, typically have an advantage in verbal ability and perceptual speed (Feingold, 1988; Torres et al., 2006; Camarata & Woodcock, 2006). In fact, Majeres (1983) found that females outperformed males in four different types of perceptual speed tasks, particularly in the WAIS coding (formerly digit symbol) task (Kaufman, McLean & Reynolds, 1988; Luciano et al., 2004).

Studying what might produce sex differences in performance in tasks related to particular areas of intelligence, Feingold (1992) examined the national data from several standardized tests (i.e., the Differential Aptitude Tests, DAT; Preliminary Scholastic Aptitude Test, PSAT; Scholastic Aptitude Test, SAT; WAIS; and California Achievement Tests, CAT). He found that males consistently had more variability in performance than females in areas such as general knowledge, mechanical reasoning, quantitative ability, spatial visualization and spelling. Feingold's finding implies that, for these certain areas of intelligence, the distribution of performance scores may be narrower for women than men. In other words, more women than men would perform near the average level, whereas men's lower and upper levels would extend further than women's in particular areas of intelligence. Feingold also found one exception: there were no sex differences in the variability on tests of perceptual speed. That is, males and females tended to be equally represented at lower, average and upper levels of performance in those tests.

It is possible that neurological differences between the sexes may be related to the differences that have been found in certain areas of intelligence. Njemanze (2005), for example,

took bilateral simultaneous transcranial Doppler ultrasound images to measure mean blood flow velocities in the middle cerebral arteries of subjects while they performed Raven's Progressive Matrices, which are non-verbal abstract reasoning tasks. He found that, while performing these tasks, males mainly used the right half of the brain; females, by contrast, mainly used the left half of the brain.

Research into sex differences in response time (RT) provides some evidence that males and females use different strategies when performing choice-RT tasks (Adam et al., 1999; Welsh & Elliott, 2001). In Adam et al., for example, participants completed a choice-RT task that required them to vocalize the position of an "X" flashed on the computer screen (e.g., "one" if the stimulus was on the left side of the screen, "two" if it was on the right). Number of choices (either two or four) and naming of the position (either compatible or incompatible) were both manipulated. They found that males had shorter RTs in all conditions. Furthermore, males seemed to employ a dichotomizing strategy (i.e., in the four-choice condition, they broke the screen into left and right parts and then chose an answer), whereas the females used a serial processing strategy (i.e., in the four-choice condition, they examined the screen from left to right and then chose an answer). Additionally, Welsh and Elliott found that females were likely trading speed for accuracy in a dichotic listening task, accounting for males' overall faster response time. This apparent male superiority has been found both in choice-RT tasks (Der & Deary, 2006; Pesta, Bertsch, Poznanski & Bommer, 2008) and in simple-RT tasks (Der & Deary, 2006).

Males have also been shown to perform more consistently than females on certain RT tasks. Roalf, Lowery & Turetsky (2006) used RT tasks in which a global stimulus was made up of local stimuli. For example, a large letter (e.g. 'H') was made up of smaller letters (e.g. 'E'), or

vice versa. Participants were instructed to press a button every time they perceived the letter 'H'. Across the global- and local-target RT tasks, males performed equally fast, whereas females performed significantly slower on the global-target tasks. Roalf et al. also demonstrated sex differences in brain activation when responding to such tasks, showing that these differences can exist on a neurological level.

Moreover, males tend to perform faster than females in RT tasks that require more physical effort. In Spierer, Petersen, Duffy, Corcoran, and Rawls-Martin (2010), for example, participants began the task with their right foot on a sensor. Using either visual (i.e., the word "GO" flashed on a television screen in front of them) or auditory stimuli (i.e., the word "GO" was heard through a speaker system), participants had to sprint a few steps to another sensor and hit it with their left foot. Researchers found male athletes to have faster RTs than females in the auditory condition, as well as faster movement times than female athletes to both visual and auditory stimuli.

In summary, several studies have shown that males have a distinct advantage in various types of RT tasks. The goal of the present study is to examine sex differences in intelligence and RT further. Specifically, the aim of this study is to test whether the sex differences in accuracy and RT to perceptual tasks are modulated by the underlying nature of the RT task. If that is the case, performance in verbal ability tests, for example, should correlate to performance in RT tasks such as lexical decision. Leite (2009) explored the relationship between IQ estimates and performance in two choice-RT tasks: brightness- and letter-discrimination tasks. He found that verbal ability rather than processing speed was more strongly correlated with mean RT in both choice-RT tasks, but this correlation was only significant in the letter discrimination task. The

current study adds two RT tasks and two intelligence subtest scores to the design in Leite in order to test task-specificity effects across gender.

The first part of the present study focused on sex differences in perceptual speed tasks and verbal ability. Perceptual speed was measured using the Symbol Search and Coding subtests of the WAIS-IV (Wechsler, 2008); verbal ability was measured by the Similarities, Vocabulary and Information subtests. Working memory and perceptual reasoning abilities were also measured by the WAIS-IV. Consistent with Camarata & Woodcock (2006), Feingold (1992), and Torres et al. (2006), we hypothesized that female scores would be higher than male scores for both perceptual speed and verbal ability. The second part of the study focused on sex differences in RT tasks. Participants completed a series of tasks involving one simple-RT (i.e., signal detection) and three choice-RT tasks. Choice-RT tasks varied in complexity based on the nature of the choice, including brightness discrimination, letter discrimination, and lexical decision. Consistent with Landauer, Armstrong & Digwood (1980), we hypothesized that males would outperform females on the simplest task, namely, the signal detection task. Assuming that verbal ability would translate into success in lexical decisions, we hypothesized that female participants would outperform males in the most complex task, i.e. the lexical decision task. For the two intermediate tasks, i.e., the brightness discrimination and the letter discrimination tasks, we hypothesized there would be no sex differences because females' advantage over males in perceptual speed might offset their disadvantage in movement speed (a motor measurement of physical speed).

### **Method**

Following Leite (2009), we used the WAIS-IV to estimate IQ (via four indices in different areas of intelligence) and four tasks to measure RT and accuracy. We used four RT



tasks with varying levels of complexity such that potential differences in performance could be examined in light of the scores in the four intelligence subareas.

Participants attended two separate sessions on different days. The first session consisted of an IQ score estimation using the WAIS-IV. In the second session, participants completed four response time (RT) tasks, namely, signal detection, brightness discrimination, letter discrimination, and lexical decision. In the latter three tasks (i.e., choice-RT tasks), masking was used in half the trials. In each RT task, participants were prompted to be as fast and as accurate as possible.

### *Participants*

Thirty-five Ohio State University students participated in the study, for which they received credit in their introductory psychology course for up to four hours of experimental participation. Five participants elected not to complete the study, and data from one participant were not analyzed because they showed around-chance performance in the choice-RT tasks. Out of the 29 participants whose data were analyzed (15 male), average age was 19 years 2 months for females (standard deviation 7.55 months) and 19 years 11 months for males (standard deviation 27.68 months).

### *Materials*

The Wechsler Adult Intelligence Scale IV (WAIS-IV; Wechsler, 2008) was used to estimate the IQ of each participant. The ten standard subtests (Block Design, Similarities, Digit Span, Matrix Reasoning, Vocabulary, Arithmetic, Symbol Search, Visual Puzzles, Information, and Coding) were used, deriving the corresponding scores in four index scales: Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI), and Processing Speed (PSI), as well as a full-scale intelligence quotient (FSIQ).

RT tasks were administered on computers running a real-time system over the Linux operating system, connected to a 17" monitor with a resolution of 640 x 480 pixels and a standard 102-key keyboard. These tasks were as follows:

*Signal Detection.* In this task, participants were asked to press the "/" key as soon as they saw a square appear on the computer screen. This was the simplest RT task used in the study, which most closely measured movement time. The squares were randomly placed in the 640 by 480-pixel screen and varied in size (202 or 502 pixels) across trial (see Figure 1 for an illustration of a trial). The task consisted of 16 blocks of 44 trials after a practice block of 20 trials.

*Brightness Discrimination.* In this task, participants were shown 160x100-pixel stimuli, randomly filled with black and white dots against a gray background (Figure 2). Their task was to judge whether each stimulus appeared bright (i.e. had more white dots) or dark (i.e. had more black dots). Stimuli were shown for 100 ms, masked by a checkered grid in half the trials. Response keys were "Z" and "/"; for a random half the participants, they corresponded to "brighter" and "darker", whereas for the other half they corresponded to "darker" and "brighter", respectively. Two levels of contrast, representing stimulus difficulty, were used: 55% of bright (or dark) pixels (easy) and 52.5% of bright (or dark) pixels (difficult). The task consisted of 16 blocks of 44 trials each, preceded by 20 practice trials (see Figure 3 for a trial illustration).

*Letter Discrimination.* In this task, participants were shown one letter drawn from one of two possible pairs (viz., F and G or B and R) and asked to indicate which letter they saw by pressing the corresponding key. Response keys were "Z" and "/"; for a random half the participants, the "Z" key was matched with "F" and "B", whereas for the other half the "/" key was matched with "F" and "B". The two pairs of letters represented easy discrimination (i.e., F

and G) and difficult discrimination (i.e., B and R), according to Gilmore, Hersh, Caramazza, and Griffin (1979). The letters were approximately 2.7 degrees in height and were shown in the center of the screen for 50 ms, followed by a mask in half of the trials. The task consisted of 18 blocks of 50 trials each, preceded by 20 practice trials (see Figure 4 for a trial illustration).

*Lexical Decision.* In this task, participants were presented with a series of letters that either formed actual words in the English language or not. They were asked to indicate whether the letters formed an actual word or not by pressing either the “Z” key or the “/” key. A random half of the participants was asked to depress the “Z” key for words and the “/” key for non-words, whereas for the other responded according to the reverse mapping. The string of letters was approximately 2.7 degrees in height and was shown in the center of the screen for 200 ms. In half the trials, the string was then masked by a row of asterisks (\*\*\*\*\*). The task consisted of 18 blocks of 40 trials each, preceded by 20 practice trials (see Figure 5 for a trial illustration).

### *Procedure*

First, each participant was separately administered the ten subtests from the WAIS-IV. This session took approximately one hour and twenty minutes to one hour and forty-five minutes. In a second session on a different day, participants were administered the RT tasks, which lasted for approximately one hour and forty-five minutes, with a ten-minute break at the midway point. Task order was randomized. Approximately half of the participants first completed the letter discrimination and lexical decision tasks, whereas the other half completed the brightness discrimination and signal detection tasks. After the break, the participants completed the remaining block of tasks.

## Results

Table 1 shows mean IQ scores and subscores for males and females. Mixed Model Analysis of Variance (ANOVA) showed no main effect of sex in the comparison of VCI, PRI, WMI, PSI and FSIQ means:  $F(1,27) = 0.05$ ,  $MSE = 86.32$ , ns;  $F(1,27) = 0.10$ ,  $MSE = 120.53$ , ns;  $F(1,27) = 3.05$ ,  $MSE = 102.67$ , ns;  $F(1,27) = 1.77$ ,  $MSE = 211.50$ , ns;  $F(1,27) = 0.00$ ,  $MSE = 79.98$ , ns, respectively.

Table 2 and Figure 6 show mean RTs for males and females on each task. In the signal detection task, an ANOVA showed significant main effects of sex and difficulty:  $F(1,27) = 4.59$ ,  $MSE = 4746.74$ ,  $p = 0.04$ ;  $F(1,27) = 32.08$ ,  $MSE = 51.74$ ,  $p < 0.01$ , respectively. Cohen's effect sizes for the sex differences in both difficulty conditions were minimal:  $d = 0.016$ . Thus, although there was a significant difference between genders in this task, the difference may not be practically important. There was no significant interaction among factors:  $F(1,27) = 0.05$ , ns.

Mean RTs in the difficult and easy conditions in the signal detection task correlated strongly with each other ( $r = 0.98$ ,  $p < 0.01$ ), indicating that performance on one level of difficulty almost perfectly predicted performance on the other level. Both difficult- and easy-condition mean RTs also correlated moderately strong with FSIQ ( $r = -0.41$ ,  $p = 0.03$ ;  $r = -0.45$ ,  $p = 0.01$ , respectively); mean RT in the easy condition correlated moderately with PRI ( $r = -0.40$ ,  $p = 0.03$ ).

Because the ANOVA tests revealed a sex difference in RTs, correlations between WAIS-IV scores and RTs were also calculated separately for males and females. Males correlated moderately with VCI in both the easy and difficult conditions ( $r = -0.52$ ,  $p = 0.05$ ;  $r = -0.56$ ,  $p = 0.03$ , respectively). Females correlated moderately with WMI in both the easy and difficult

conditions, ( $r = -0.53$ ,  $p = 0.05$ ;  $r = -0.50$ , ns, respectively) although the correlation was significant only in the easy condition.

In the brightness discrimination task, an ANOVA found no significant sex difference in the RT means:  $F(1,27) = 0.61$ ,  $MSE = 41653.82$ , ns. There were, however, significant main effects of masking and difficulty:  $F(1,27) = 12.77$ ,  $MSE = 7069.10$ ,  $p = 0.01$ ;  $F(1,27) = 8.36$ ,  $MSE = 558.79$ ,  $p < 0.01$ , respectively. Participants performed more quickly in the easy condition rather than the difficult condition and surprisingly, in the masked condition rather than the non-masked condition. (This masking advantage may have occurred since the mask used was a checkered grid with 50% of both white and black squares. Thus, participants may have compared the stimulus to the mask in order to determine whether the stimulus had a higher dark to bright or bright to dark ratio).

An ANOVA on task accuracy found no significant sex differences:  $F(1,27) = 1.58$ ,  $MSE = 0.06$ , ns. There was also no significant difference in accuracy between masked and unmasked conditions:  $F(1,27) = 0.02$ ,  $MSE = 0.02$ , ns. However, there was a significant main effect in accuracy of difficulty levels:  $F(1,27) = 168.91$ ,  $MSE = 0.00$ ,  $p < 0.01$ . Participants were much more accurate in the easy condition than in the difficult condition. An interaction between mask and difficulty was also found to be significant:  $F(1,27) = 20.04$ ,  $MSE = 0.00$ ,  $p < 0.01$ ; indicating that the difference between difficulty conditions was smaller in the masked condition. No other interactions were significant. Table 3 displays the mean accuracy for males and females on each task.

The brightness discrimination mean RTs correlated strongly with each other ( $r = 0.69$  to  $0.97$ ,  $p < 0.01$ ). Correlations between all difficulty levels and masking conditions were

significant. Furthermore, the easy non-masked, difficult non-masked and difficult masked condition means moderately correlated with PRI ( $r = -0.43, p = 0.02$ ;  $r = -0.46, p = 0.01$ ;  $r = -0.39, p = 0.04$ , respectively). Additionally, the difficult, non-masked brightness discrimination task mean correlated moderately with PSI ( $r = -0.39, p = 0.04$ ). With regards to accuracy, the brightness discrimination tasks correlated strongly with each other ( $r = 0.61$  to  $0.97, p < 0.01$ ).

Correlations were also calculated between task RT means and respective accuracies to determine if a speed-accuracy trade-off existed. Females' correlation between mean RT and accuracy was significant in the easy non-masked, easy masked and difficult masked conditions ( $r = 0.56, p = 0.04$ ;  $r = 0.59, p = 0.03$ ;  $r = 0.68, p < 0.01$ , respectively). Males' correlation between mean RT and accuracy was significant in only the difficult non-masked condition ( $r = 0.52, p = 0.05$ ).

In the letter discrimination task, an ANOVA found no significant sex difference in the RT means:  $F(1,27) = 0.02, MSE = 17963.73, ns$ . There was also no difference between masked and unmasked condition RT means:  $F(1,27) = 1.88, MSE = 3691.43, ns$ . However, a difference was found between difficulty condition RT means:  $F(1,27) = 6.75, MSE = 318.93, p = 0.02$ . Participants were faster to distinguish between F and G than between B and R. No main interactions were found to be significant. An ANOVA on task accuracy found no significant sex differences:  $F(1,27) = 1.85, MSE = 0.04, ns$ . There was, though, a significant difference in accuracy between masked and unmasked conditions:  $F(1,27) = 28.28, MSE = 0.01, p < 0.01$ . Participants were more accurate on unmasked trials than on masked trials. There was no main effect in accuracy of difficulty:  $F(1,27) = 2.23, MSE = 0.00, ns$ .

The letter discrimination mean RTs correlated strongly with each other ( $r = 0.62$  to  $0.95, p < 0.01$ ). Correlations between all difficulty levels and masking conditions were significant.

Both the easy non-masked and difficult non-masked condition means also moderately correlated with FSIQ ( $r = -0.37, p = 0.05$ ;  $r = -0.39, p = 0.03$ , respectively). Furthermore, the difficult non-masked and easy masked condition means moderately correlated with PRI ( $r = -0.38, p = 0.04$ ;  $r = -0.37, p = 0.05$ ). With regards to accuracy, the letter discrimination tasks correlated strongly with each other ( $r = 0.57$  to  $0.93, p < 0.01$ ).

Females had no significant correlations between mean RTs and accuracy in any of the conditions, indicating that overall, neither RT nor accuracy were predictive of each other. Males' correlations between mean RT and accuracy were strong in the easy and difficult non-masked conditions ( $r = -0.68, p < 0.01$ ;  $r = -0.73, p < 0.01$ ). These negative correlations indicate that for males there was no speed-accuracy trade-off: they completed the task both quickly and accurately.

In the lexical decision task, an ANOVA found no significant sex difference in RT means:  $F(1,27) = 0.09, MSE = 57415.97, ns$ . There were, however, significant main effects of masking, word frequency and word legitimacy:  $F(1,27) = 9.81, MSE = 2015.99, p < 0.01$ ;  $F(1,27) = 107.37, MSE = 776.85, p < 0.01$ ;  $F(1,27) = 48.23, MSE = 2426.99, p < 0.01$ , respectively. Participants were faster in the masked condition than the non-masked, in the high-frequency rather than the low-frequency condition, and in the word rather than the non-word condition. An interaction between mask and frequency was significant:  $F(1,27) = 4.66, MSE = 717.19, p = 0.04$ ; indicating that there was no difference between masked and unmasked condition RTs in the high-frequency word condition. An interaction between frequency and legitimacy was also significant:  $F(1,27) = 84.14, MSE = 1226.98, p < 0.01$ ; indicating no difference between the word & non-word conditions in the low-frequency condition. Finally, an interaction between

mask, frequency and legitimacy was significant:  $F(1,27) = 5.86$ ,  $MSE = 395.10$ ,  $p = 0.02$ . No other mean interactions were found to be significant.

An ANOVA on task accuracy found no significant sex differences:  $F(1,27) = 1.75$ ,  $MSE = 0.03$ , ns. However, a main effect of word frequency in accuracy was significant:  $F(1,27) = 101.62$ ,  $MSE = 0.00$ ,  $p < 0.01$ ; indicating that participants were more accurate at identifying high-frequency words than low-frequency words. There was no significant difference in accuracy between word and non-word conditions:  $F(1,27) = 0.68$ ,  $MSE = 0.01$ , ns. An interaction between accuracy in frequency and word legitimacy was significant:  $F(1,27) = 128.84$ ,  $MSE = 0.00$ ,  $p < 0.01$ ; indicating that there was no difference between high- and low-frequency word accuracy in the non-word condition. Additionally, an interaction between mask, frequency and legitimacy was found:  $F(1,27) = 5.07$ ,  $MSE = 0.00$ ,  $p = 0.03$ . No other interactions were significant.

Furthermore, the lexical decision task means correlated strongly with each other ( $r = 0.70$  to  $0.97$ ,  $p < 0.01$ ). This indicates that good performance on one level of difficulty, or being masked vs. non-masked, or being a word vs. non-word, predicted performance on other levels. Each condition mean significantly correlated strongly with FSIQ ( $r$  ranging from  $-0.42$  to  $-0.61$ ,  $p < 0.01$ ). Each condition mean also correlated with PRI ( $r$  ranging from  $-0.39$ ,  $p = 0.04$  to  $-0.64$ ,  $p < 0.01$ ). Furthermore, every condition mean RT except for high frequency, no-masked words, correlated strongly with VCI ( $r$  ranging from  $-0.45$ ,  $p = 0.01$  to  $-0.60$ ,  $p < 0.01$ ). The mean RT for high frequency, no-masked word means was trending towards significant ( $r = -0.36$ ,  $p = 0.054$ ). With regards to accuracy, the lexical decision tasks correlated significantly with each other ( $r = 0.38$ ,  $p = 0.04$  to  $0.74$ ,  $p < 0.01$ ). Additionally, the accuracy in low-frequency masked



and non-masked conditions strongly correlated with VCI ( $r = 0.57, p < 0.01$ ;  $r = 0.48, p < 0.01$ , respectively).

Females' correlation between mean RT and accuracy was strong in the non-masked low frequency word and non-masked high frequency non-word conditions ( $r = -0.60, p = 0.02$ ;  $r = -0.62, p = 0.02$ , respectively). Males' correlation between mean RT and accuracy was moderately strong in the non-masked low frequency non-word, non-masked high frequency non-word and masked high-frequency non-word conditions ( $r = -0.54, p = 0.04$ ;  $r = -0.54, p = 0.04$ ;  $r = -0.59, p = 0.02$ , respectively). Again, these negative correlations indicate that there was no trade-off between speed and accuracy.

### **Discussion**

This study examined whether sex differences exist in specific areas of intelligence and in RT tasks of different complexity levels. In accordance with our hypothesis (cf., Landauer, Armstrong & Digwood, 1980; Der & Deary, 2006), we found that males outperformed females on the signal detection task. We attributed this result to males' advantage in movement time rather than any type of advantage in perceptual speed because males tend to have more muscle fiber, which allows them to perform physical actions more quickly than females (Spierer, Petersen, Duffy, Corcoran & Rawls-Martin, 2010).

Contrary to what we hypothesized, we did not find sex differences in any of the WAIS-IV subtests or choice-RT tasks. It is possible that we had too small of a sample size to identify these differences. Previously reported sex differences in these areas have been typically small and found in larger samples. In addition, contrary to previous studies that have reported a female advantage (e.g., Majeres, 1983 and Torres et al., 2006), we found no sex differences in the PSI or

VCI subscores, which may help explain the lack of sex differences in the choice-RT tasks. Thus, it is possible that our hypotheses for a female advantage in the lexical decision RT task could be supported by a larger sample size in which a female advantage was found in PSI and VCI.

Although we did not find sex differences in the choice-RT tasks, the gender-specific correlations between mean RT and intelligence subareas found in the signal detection task may reflect how males and females use different resources to perform the same simple task. Males correlated more strongly with VCI whereas females correlated more strongly with WMI. Perhaps this means that males use some sort of verbal component whereas females use a working memory component to perform this task. How verbal ability or working memory could help one perform a simple perceptual task is still open to further study.

In studying gender differences in choice-RT tasks, however, it is natural to question whether speed-accuracy trade-offs confounded or masked such differences. We found evidence of speed-accuracy trade-offs in the brightness discrimination task only, in which females, but not males, tended to slow down to achieve higher accuracy. These results are consistent with Welsh & Elliott's finding (2001) that females were likely trading speed for accuracy, but they were not strong enough in our case to drive gender differences in mean RT or accuracy. In the letter discrimination task, there was no indication that females were trading speed for accuracy, and males were able to perform the task both quickly and accurately. In the lexical decision task, both males and females were able to complete the tasks quickly and accurately. It is possible that the speed-accuracy trade-off was only observed in the brightness discrimination task because it was the only task in our study that was challenging for participants. From that standpoint, our results point to a more zealous performance by females than by males in face of a more challenging task.

In addition, evidence from this study suggests that the choice-RT tasks we used are not perfectly analogous to tasks used to measure PSI. The previously found female advantage in the coding subtest of the WAIS, which measures PSI, did not correlate with any advantages in choice-RT tasks. Furthermore, PSI correlated only with one condition in the brightness discrimination task. Surprisingly, almost every condition on every task correlated with PRI. Also, PSI and PRI were also the only two WAIS-IV subtests that significantly correlated with each other. These findings suggest the PRI subtests on the WAIS-IV may have been capturing a PSI component or conversely, that the RT tasks we used were unintentionally measuring a component of PRI. Sex differences may indeed exist in particular RT tasks, but the trouble lies in finding which tasks most cleanly capture PSI and thus will show biases towards one gender or the other.

In summary, the findings from this study showed that the previously found female advantage in perceptual speed and verbal ability may be undermined by other abilities (or skills) that, jointly, are recruited for perceptual decision making. Furthermore, we found that males and females possibly employed different strategies when completing the signal detection task (their response times correlated differently with the four WAIS-IV subscores), consistent with Adam et al., 1999; Welsh & Elliott, 2001; and Njemanze, 2005, who found that males and females used different resources when completing various tasks.

### References

- Adam, J. J., Paas, F. G. W. C., Buekers, M. J., Wuyts, I. J., Spijkers, W. A. C., & Wallmeyer, P. (1999). Gender differences in choice reaction time: Evidence for differential strategies. *Ergonomics*, 42(2), 327-335. doi:10.1080/001401399185685
- Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. *Intelligence*, 34(3), 231-252. doi:10.1016/j.intell.2005.12.001
- Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, 21(1), 62-73. doi:10.1037/0882-7974.21.1.62
- Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist*, 43(2), 95-103.
- Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research*, 62, 61-84.
- Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics*, 25, 425-431.
- Gras, R., Bordoy, M., Ballesta, G., & Berna, J. (2010). Creativity, intellectual abilities and response styles: Implications for academic performance in the secondary school. *ANALES DE PSICOLOGIA*, 26(2), 212-219.
- Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12(2), 229-246. doi:10.1023/A:1009027516424

- Johnson, W., Bouchard, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, 32(1), 95-107.  
doi:10.1016/S0160-2896(03)00062-X
- Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS—R subtests. *Journal of Clinical Psychology*, 44(2), 231-248. doi:10.1002/1097-4679(198803)44:2<231::AID-JCLP2270440224>3.0.CO;2-J
- Landauer, A.A., Armstrong, S. & Digwood, J. 1980, Sex difference in choice reaction time, *British Journal of Psychology*, 71, 551-555.
- Leite, F. (2009). Should IQ, perceptual speed, or both be used to explain response time? *American Journal of Psychology*, 4, 517-526.
- Luciano, M., Wright, M., Geffen, G., Geffen, L., Smith, G. & Martin, N. (2004). A genetic investigation of the covariation among inspection time, choice reaction time, and IQ subtest scores. *Behavior Genetics*, 34(1), 41-50.
- Lynn, R. (1998). Sex differences in intelligence: Data from a Scottish standardisation of the WAIS-R. *Personality and Individual Differences*, 24(2), 289-290. doi:10.1016/S0191-8869(97)00165-7
- Majeres, R.L. (1983). Sex differences in symbol-digit substitution & speeded matching. *Intelligence*, 7, 313-327.
- Njemanze, P. C. (2005). Cerebral lateralization and general intelligence: Gender differences in a transcranial Doppler study. *Brain and Language*, 92(3), 234-239.
- Nyborg, H. (2005). Sex-related differences in general intelligence g, brain size, and social status. *Personality and Individual Differences*, 39(3), 497-509. doi:10.1016/j.paid.2004.12.011

- Pesta, B. J., Bertsch, S., Poznanski, P. J., & Bommer, W. H. (2008). Sex differences on elementary cognitive tasks despite no differences on the Wonderlic Personnel Test. *Personality and Individual Differences*, 45(5), 429-431. doi:10.1016/j.paid.2008.05.028
- Roalf, D., Lowery, N., & Turetsky, B. I. (2006). Behavioral and physiological findings of gender differences in global-local visual processing. *Brain and Cognition*, 60(1), 32-42. doi:10.1016/j.bandc.2005.09.008
- Silverman, I. (2006). Sex differences in simple visual reaction time: A historical meta-analysis. *Sex Roles*, 54(1-2), 57-68. doi:10.1007/s11199-006-8869-6
- Spierer, D., Petersen, R., Duffy, K., Corcoran, B. & Rawls-Martin, T. (2010). Gender influence on response time to sensory stimuli. *Journal of Strength and Conditioning Research*, 24(4), 957-963. doi:10.1519/JSC.0b013e3181c7c536
- Torres, A., Gomez-Gil, E., Vidal, A., Puig, O., Boget, T., & Salamero, M. (2006). Gender differences in cognitive functions and influence of sex hormones. *Actas Espanolas de Psiquiatria*, 34(6), 408-415.
- Wechsler, D. (1981). The psychometric tradition – Developing the Wechsler Adult Intelligence Scale. *Contemporary Educational Psychology*, 6(2), 82-85.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition*. San Antonio, TX: Pearson Assessment.
- Welsh, T. N. & Elliott, D. (2001). Gender differences in a dichotic listening and movement task: Lateralization or strategy? *Neuropsychologia*, 39(1), 25-35. doi:10.1016/S0028-3932(00)00094-4

Table 1

*Mean WAIS-IV Scores*

	<b>N</b>	<b>Mean</b>	<b>Std Dev</b>	<b>Minimum</b>	<b>Maximum</b>
VCI	29	92.55	9.13	63	108
Male	15	92.93	9.00	78	108
Female	14	92.14	9.60	63	98
PRI	29	90.10	10.80	60	109
Male	15	90.73	11.67	60	109
Female	14	89.43	10.18	69	109
WMI	29	91.76	10.50	71	108
Male	15	94.93	8.13	83	108
Female	14	88.36	11.92	71	105
PSI	29	108.28	14.74	79	150
Male	15	104.80	12.83	79	122
Female	14	112.00	16.19	81	150
FSIQ	29	93.66	8.39	77	108
Male	15	93.73	9.57	77	108
Female	14	93.57	7.27	84	108

Table 2

*Mean RT for Each Task*

	Sig Mean RT		Bri Mean RT			
	Easy	Diff	Easy Non-Mask	Diff Non-Mask	Easy Mask	Diff Mask
Male	279	290	599	609	546	565
Female	319	328	576	587	514	524

	Let Mean RT				Lxc Mean RT			
	Easy Non-Mask	Diff Non-Mask	Easy Mask	Diff Mask	Low Non-Mask	High Non-Mask	Low Mask	High Mask
Male	516	526	529	543	664	561	635	568
Female	510	525	536	531	647	562	621	555

*Note.* Response times are in milliseconds.



Table 3

*Mean Accuracy*

	Bri Accuracy				Let Accuracy			
	Easy Non-Mask	Diff Non-Mask	Easy Mask	Diff Mask	Easy Non-Mask	Diff Non-Mask	Easy Mask	Diff Mask
Male	0.832	0.702	0.783	0.736	0.904	0.893	0.799	0.798
Female	0.769	0.639	0.738	0.671	0.881	0.854	0.730	0.717

	Lxc Accuracy			
	Low Non-Mask	High Non-Mask	Low Mask	High Mask
Male	0.824	0.958	0.797	0.951
Female	0.786	0.922	0.765	0.915

*Note.* Accuracies are listed in percentages. Accuracy was not a dependent measure in the signal detection task.

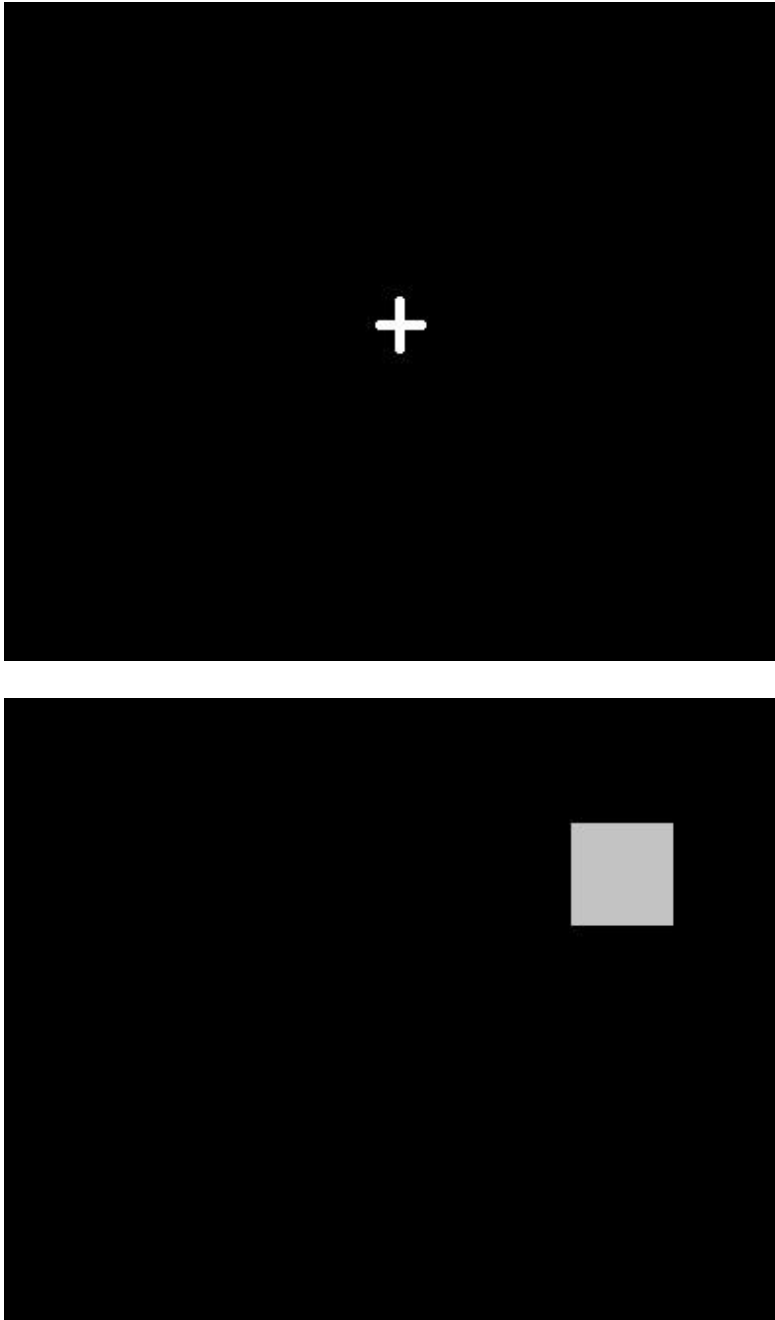


Figure 1. An illustration of the signal detection task. The top picture represents the fixation screen; the bottom picture is an example of a stimulus.

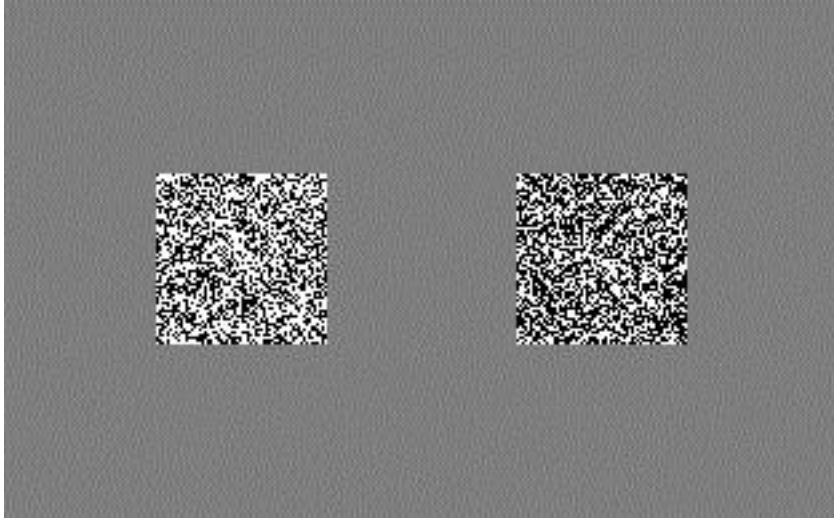


Figure 2. Examples of easy stimuli in the brightness discrimination task. A “brighter” stimulus is shown on the left; a “darker” stimulus is shown on the right.

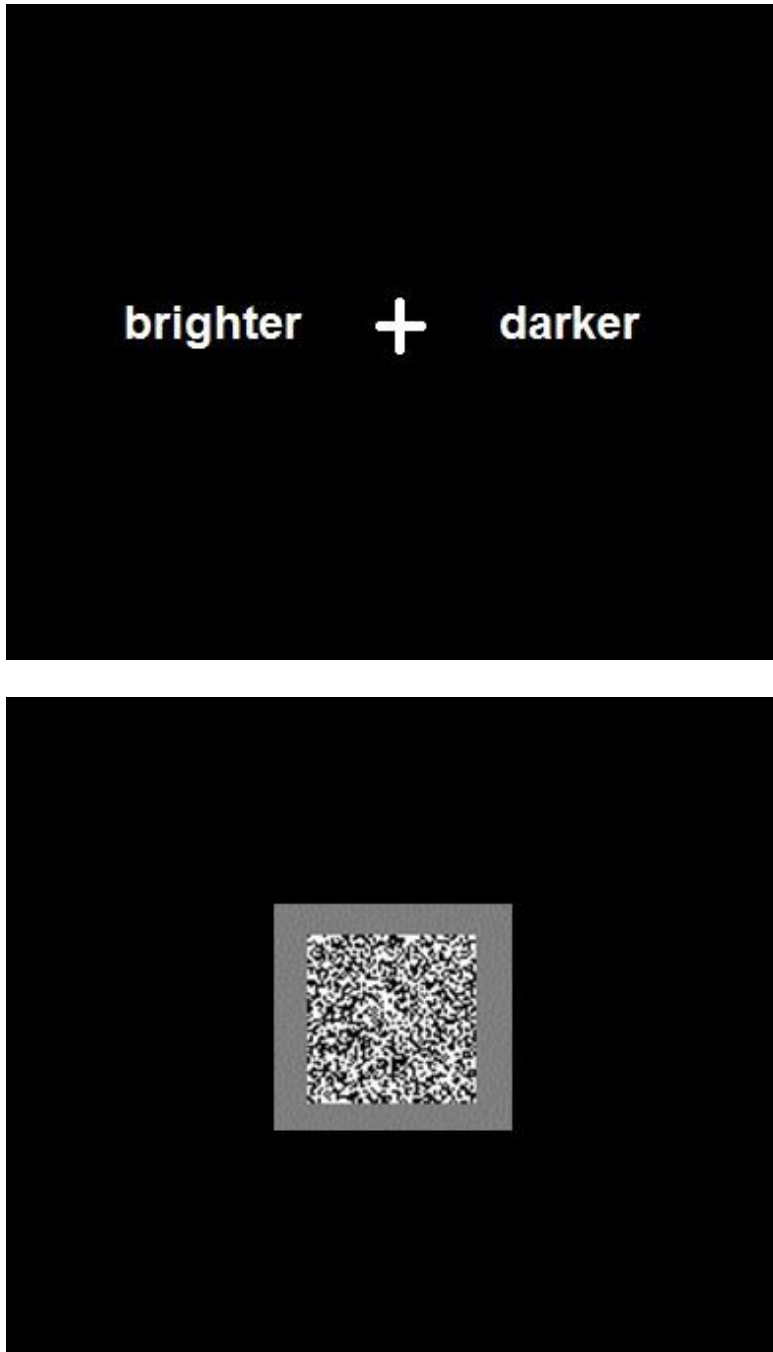


Figure 3. An illustration of the brightness discrimination task. The top picture represents the fixation screen; the bottom picture is an example of a stimulus.

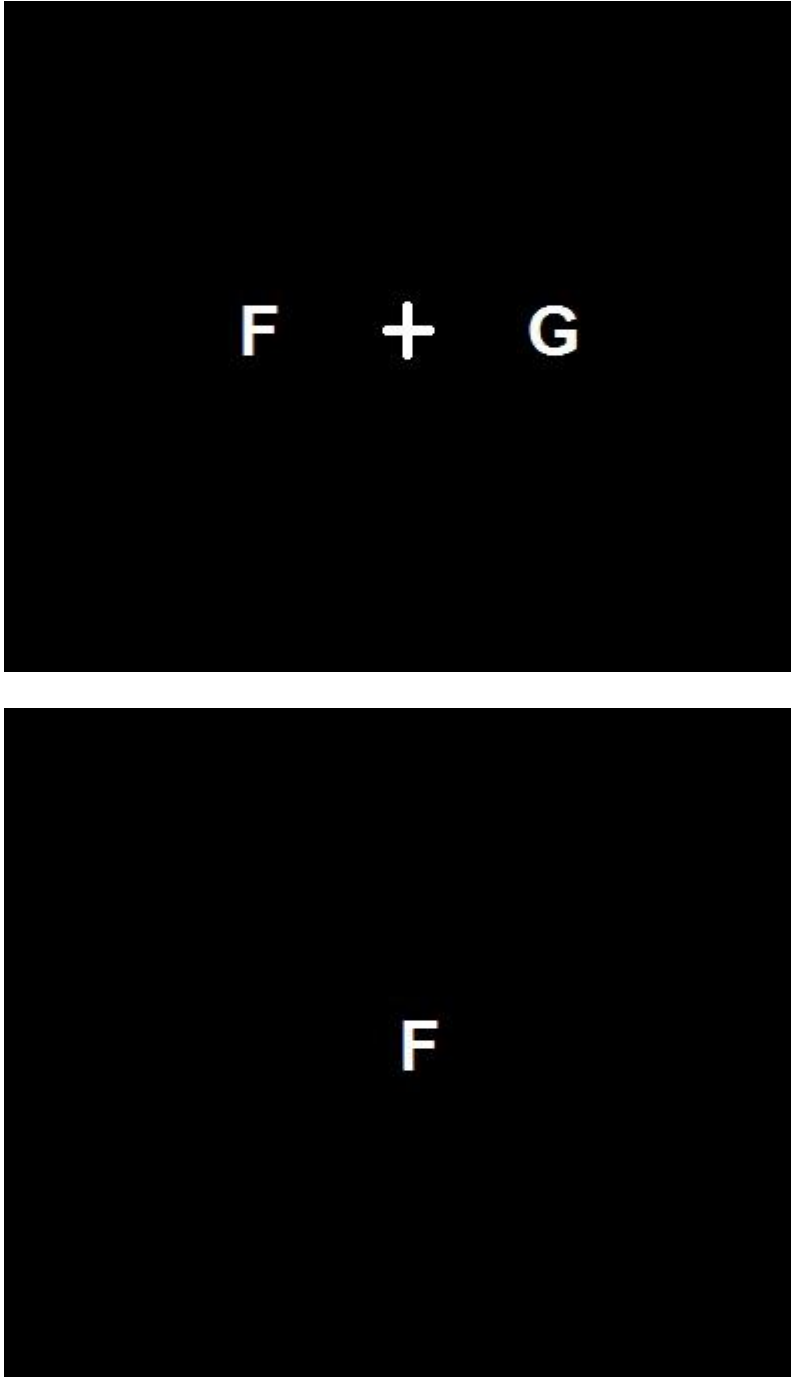


Figure 4. An illustration of the letter discrimination task. The top picture represents the fixation screen; the bottom picture is an example of a stimulus.

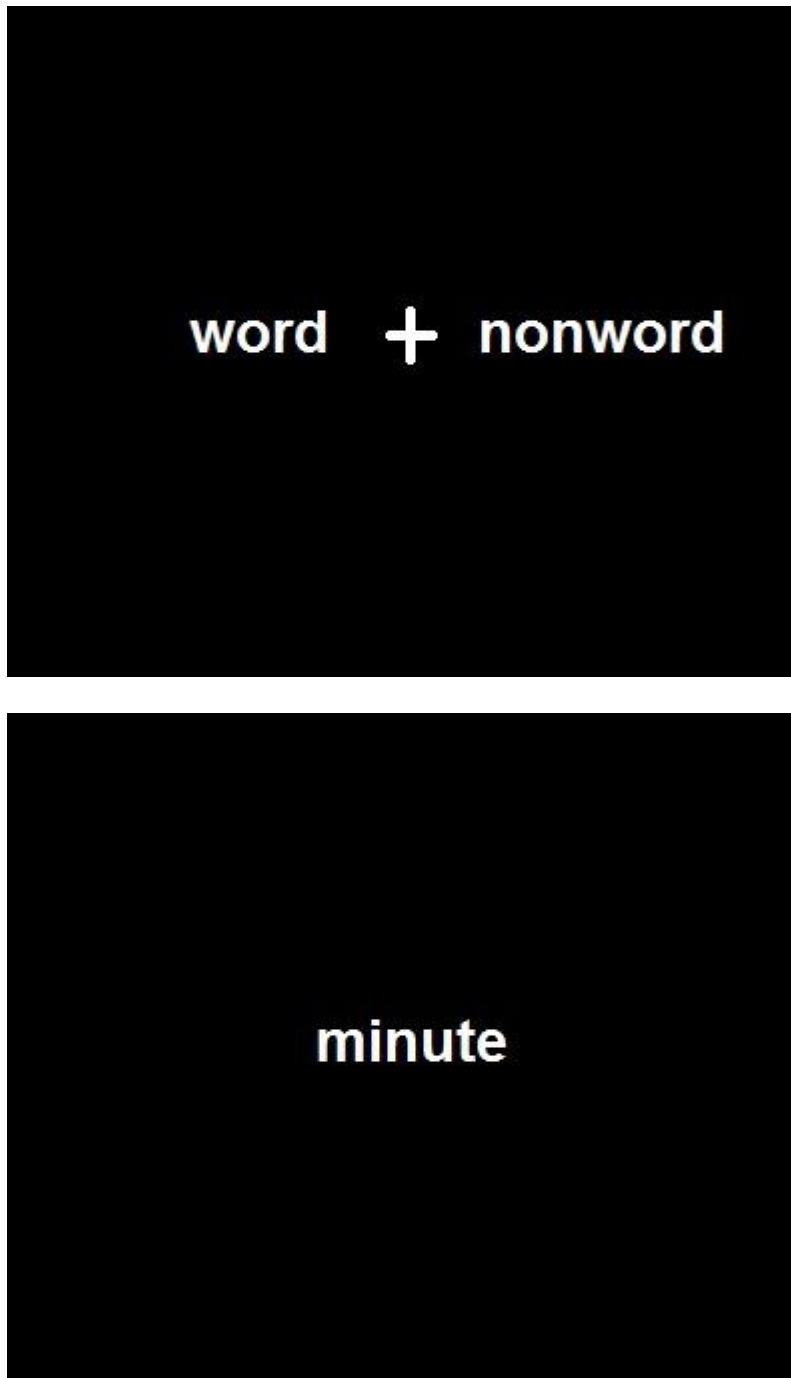


Figure 5. An illustration of the lexical decision task. The top picture represents the fixation screen; the bottom picture is an example of a stimulus.

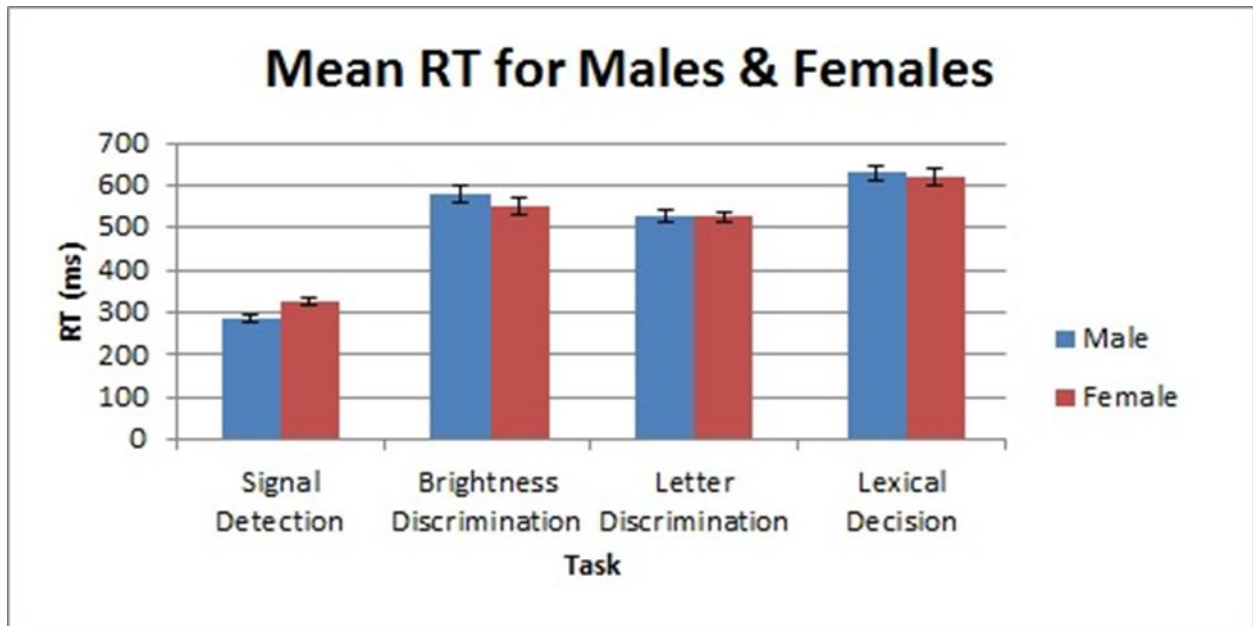


Figure 6. Mean response time across male and female participants for each of the four tasks.

Error bars indicate one standard error above and below the mean.