The Yale Review of Undergraduate Research in Psychology

A Student Publication

Fourth Issue

About

The Yale Review of Undergraduate Research in Psychology is an annual journal that showcases the best and most original research in psychology conducted by undergraduates from around the world. Our goal is to contribute to scientific advancement by encouraging serious, quality research early on. We provide a platform for undergraduate scientists to share their findings, and aim to bring together a community of young psychologists from both the United States and abroad.

Acknowledgements

Staff

Contents

Mirror Writing and Hand Dominance in Children: A New Perspective on Motor and Perceptual Theories

Ailbhe Brennan
The University of Edinburgh

Mirror writing in nursery and school-aged children was investigated using a novel approach. The motor hypothesis of mirror writing, which proposes that the non-dominant hand may be more adept at mirror writing, was assessed with dominant and non-dominant hand writing using a digitizing tablet. A measure of perceptual discrimination was used to investigate the relationship between mirror writing and the perceptual hypothesis, which states that unintentional mirror writing may be attributable to a form of perceptual confusion. Findings demonstrated a significant positive correlation between mirror writing and perceptual confusion, indicating that perception is the predominant driving factor in the majority of young children. Mirror writing was shown to decrease with age and as children grow older, motor factors gradually take over as a foundation for mirror writing. As children mature, these motor mechanisms gradually become governed by cognitive control strategies. In certain cases, brain damage may occur that disrupts these control strategies, producing unintentional mirror writing in adults. This developmental perspective of mirror writing considers it to be an inherent phenomenon and a normal part of writing development in children.

Keywords: mirror writing, children, perceptual, motor, dominant, non-dominant

Mirror writing is defined as the production of individual letters or whole words in reversed form, such that they become easily legible when viewed with a mirror. Mirror writing is most commonly characteristic of young children who are in the early stages of language acquisition (Schott, 2007), and has also been identified in certain cases of brain damage in adults (Paradowski & Ginzburg, 1971). Despite there being a few famous cases of adults practicing mirror writing intentionally (Schott, 1999), for most individuals it constitutes a complex, unnatural and cognitively demanding task. In light of this fact, it is intriguing that the vast majority of children have a tendency to mirror reverse words, letters and digits at some point during their writing development (Cornell, 1985).

Several researchers have proposed theories regarding the behavioral basis of mirror writing, with the debate centered on whether mirror writing can be considered a predominantly perceptual or motor phenomenon.

Perceptual Explanations Perceptual theories of mirror writing attribute unintentional mirror writing in children to perceptual confusions of the letters. One such theory is the mirror engram hypothesis (Orton, 1928). This theory states that visual representations (engrams) of stimuli such as letters and words are stored in the dominant hemisphere for language (usually left), while the corresponding mirrored engram is stored in the alternate hemisphere. These mirrored engrams are normally suppressed unless hemispheric dominance has yet to be established (as is the case in young children), or if damage occurs to the mechanism involved in such suppression (as may be the case in brain damaged adults). Confusion over the internal representation of letters would therefore elicit mirror writing with both the dominant and non-dominant hands. A perceptual foundation may also explain the large concurrence of mirror writing with mirror reading.

Orton's mirror engram hypothesis has found support from case studies of adults with brain damage (e.g. Gottfried, Sancar & Chatterjee, 2003; Heilman, Howell, Valenstein, & Rothi, 1980) and experimental studies with normally functioning adults (Tucha, Aschenbrenner, & Lange, 2000; Tankle & Heilman, 1983). Tankle and Heilman (1983) focused on whether the left

hand was more adept at mirror writing and investigated the basis for this superiority. Participants were asked to mirror write words and sentences with both their dominant and non-dominant hands and errors were counted as instances when the writing was not correctly mirror reversed. It was found that when writing with the dominant hand, left handers mirror wrote with fewer errors ($p < .01$) and at a faster pace ($p = .04$) than right handers. There was no difference detected between right and left handers' speed when writing in normal direction ($p > .20$). Building on the work of Tankle and Heilman (1983), Tucha et al. (2000) instructed left- and right- handed subjects to mirror write with both hands using the touch screen of a digitizing tablet, making it easier to write with the non-dominant hand. This study also stressed the role of left handed superiority in mirror writing, as left handers were found to make significantly fewer errors than right handers when writing with their dominant hand ($p < .01$).

An alternative model of mirror writing was devised by Dehaene, Nakamura, Jobert, Kuroki, Ogawa and Cohen (2010) in an effort to explain the neural substrate of spontaneous mirror writing in children. Their fMRI study with adults found the human perceptual system to be inherently dichotomous, presenting evidence of an 'unlearned' capacity for recognizing mirrored forms of writing in adults. They propose that this mechanism is still intact in children who are in the early stages of language acquisition and can thus account for the readily available mirrored representation of letters and words, which in turn produces mirror reading and writing.

There exists a relatively sparse body of literature on mirror writing in normal children. Due to the misinterpretation of mirror writing as a dysfunction in children, older studies tended to focus on learning and other developmental difficulties (Orton, 1928). Recent research has dispelled such myths and has shown there to be little or no relationship between mirror writing and learning difficulties or intelligence in children (Cubelli & Della Sala, 2009).

An early study with normally developing children used a simple technique to induce mirror writing in those aged between 3 and 14 years old (Cornell, 1985). Children were asked to write their names on a sheet of paper that was bisected by a line. When instructed to write on the left side of the line, there was insufficient space for the child to write his or her name and to do so in a correct direction would require the child to write over the line. Children aged 8 years or older (n = 54) all successfully wrote their names in a left-right direction across the line. The younger group of children, however, did not always successfully write their names across the line, and instead tended to mirror write their names in a right-left direction (n = 99). The proportion of mirror writing dropped off as a function of age, ranging from 82% of 5 year olds to 13% of 7 year olds.

By contrast, Fischer and Tazouti (2012) rationalized that the perceptual explanation of mirror writing could be split into two facets: errors in the direct perception of letters (assessed by means of a copying task) and errors in the internal representation of those letters (assessed by writing from memory). A large sample of children (n = approximately 300) aged between 5 and 6 years was recruited. Under the memory condition, mirror writing was relatively frequent (> 20%); while in the copying condition the prevalence of mirror writing was negated (< 0.5%). Referring to previous results of Fischer (2010, 2011), the authors here reiterate that in the absence of a defined direction of letters, children use their implicit knowledge of orientation of characters when writing. Research has found that those letters and digits that are mirror written most often are those that end in strokes facing leftwards (e.g. J, Z, 3). As English is a predominantly rightward facing language, both in terms of the individual letters and the overall direction of script, it is assumed that children may over-apply this 'right writing rule' (Fischer, 2011).

Motor Explanations An alternative set of explanations attributes mirror writing not to perceptual factors, but rather to motor factors. According to one motor hypothesis, as first proposed by Erlenmeyer in 1879 (as cited in Critchley, 1928), the motor sequence for writing with the dominant hand is stored in the contralateral hemisphere. It is hypothesized that when writing is undertaken with the non-dominant hand, the motor sequence must be transferred to the alternate hemisphere and becomes mirrored in the process (Noble, 1968). Another motor hypothesis is related to the popularized perception that mirror writing is the natural script of the left hander. The basis for this is that adductive movements tend to be more comfortable than abductive movements (Brown, Knauft & Rosenbaum, 1948). This would suggest that when right handers undertake writing with their left hand, it may be more natural to start from the midline and write in a right-left (adductive) direction. As a consequence of this theory, it could plausibly be predicted that left handers are better able to overcome the left-right directional bias (Tankle & Heilman, 1983).

There is a growing body of research supporting motor hypotheses (Angelilo, De Lucia, Trojano & Grossi, 2010; Rodriguez, 1991; Rodriguez, Aguilar & Gonzalez, 1989). Evidence for motor mechanisms has been observed in cases of brain damaged adults (Balfour, Borthwick, Cubelli & Della Sala, 2009; Buxbaum, Coslett, Schall, MacNally & Goldberg, 1993) and in more recent studies with children (Della Sala & Cubelli, 2007; Wang, 1992). The

motor hypothesis would be testable in younger children if they were to write with their non-dominant hand, as such age groups may lack the conscious awareness to override the basic mirrored motor output.

The term 'directional apraxia' was used by Della Sala and Cubelli (2007) to explain unintentional mirror writing. Directional apraxia refers to the unavailability of the correct direction of movement. Their investigation into mirror writing concerned both patients with left hemisphere stroke and normally developing children. The authors argued that our motor systems are inherently dichotomous and in young children, an appropriate writing direction has yet to be encoded due to inexperience with writing. Concerning adults with brain damage, the theory is that the acquired writing direction is lost due to infarction. Della Sala and Cubelli argued against a perceptual explanation of mirror writing as they failed to find a relationship between mirror writing and performance on perceptual and orientation tasks in children. This study provides an interesting perspective on mirror writing by considering the two populations alongside one another. Overall, this approach succeeds in unifying the theoretical underpinnings of mirror writing across different populations. The findings were later upheld by Cubelli and Della Sala (2009) when they tested the same sample of children. A caveat should be applied to these methods, however, in that odd-one-out picture tasks were used as a measure of perception and orientation. A more appropriate task would have been a perceptual confusion task with written stimuli, as discrimination of mirrored images and letters are different processes (Pedago, Nakamura, Cohen & Dehaene, 2011).

To the best of our knowledge, only one study has explored the motor hypothesis in children by asking them to write with their non-dominant hand (Wang, 1992). Writing with the dominant and non-dominant hand was examined in normally developing children and adults. Overall, Wang found a higher prevalence of mirror writing with the left hand (45.8%) compared to the right hand (22.2%) in preschool children (n = 72). There was a significant drop off of mirror writing with the left hand (10%) in school children (n = 40), and no mirror writing was observed with the right hand. Given that right handedness is more common than left handedness, we assume these results to be indicative of non-dominant (left) and dominant (right) hands. Wang also found a relationship between left/right spatial disorientation and mirror writing in preschool children (67.6%). Writing with the non-dominant in adults did not elicit mirror writing except in one case (n = 40). This suggests a greater effect of motor driven mechanisms in children that gradually drops off with age.

The Present Study Much research in this area has focused on mirror writing in brain-damaged adults and much current thinking stems from such work. Our study, by contrast, is concerned with the prevalence of mirror writing in children. Given the relatively small body of literature on mirror writing in children and in the absence of satisfactory contradictory evidence, it is reasonable to assume that the mechanisms driving involuntary mirror writing in brain damaged adults and young children may share a common underlying neural substrate.

In light of the evidence discussed above, we propose an investigation into dominant and non-dominant hand effects on children's mirror writing. In cases where mirror reading is reported alongside mirror writing, motor hypotheses cannot account for both. It would seem, therefore, that mirror writing is not a unitary disturbance; rather it is likely that multiple processes are at work. For this reason, we accounted for the possibility of both motor and perceptual foundations in our study. There are numerous shortcomings in the research that are rectifiable by a simple experimental approach. The method pioneered by Wang (1992) has proven to be an unusually effective method of assessing the motor hypothesis in young children. The use of a tablet (Tucha et al., 2000) will eliminate any potential confounds to holding a pen with the non-dominant hand. A more concrete approach to assessing children's perceptual abilities would be by means of a letter perceptual discrimination task, as opposed to an odd-one-out task (Della Sala & Cubelli, 2007).

To test for motor factors, we had subjects write their names and a selection of letters with both their dominant and non-dominant hands. Bimanual tasks such as these allow us to see how much conscious control children are exerting on their writing direction and are thus effective in determining whether mirror writing is due to motor factors. To test for perceptual factors, we chose to use an alphabetic directional discrimination task. This allowed us to gauge the extent to which perceptual confusions impacted children's mirror writing. Using these methods meant that we could conduct straightforward correlational analyses on the data collected, resulting in discernible relationships between motor factors, perceptual factors and age. These tests also allowed us to record a single data point for each of the letters analyzed, allowing us to test Fischer's (2011) theory of the 'right writing rule'.

We predict that perceptual factors will play an overriding role in spontaneous mirror writing in younger children, while motor influences are likely to impact increasingly as the child grows older and perceptual confusions fade. We predict that we will find consistent mirror writing of certain letters across both dominant and non-dominant hands in younger children. In older

children, we expect to observe less mirror writing with the dominant hand and more mirror writing with the non-dominant hand, in accordance with Critchley (1928). As a separate hypothesis, we expect to find more mirror writing of less common, leftward-facing letters, as predicted by Fischer (2011).

METHOD

Writing with both dominant and non-dominant hands was assessed among pre-school and school-going children in order to analyze spontaneous occurrences of mirror reversals. These instances were recorded alongside literacy and perceptual measures as a means to determining the possible underlying causes of mirror writing in children.

Participant Characteristics The sample consisted of 51 normally developing children (28 boys, 23 girls; aged 48 - 124 months, mean 79.33, SD 18.38) recruited from local nurseries and after-school clubs. Only 5 children were considered to be left-handed and they were all male. Children were selected to participate on the basis that they could spontaneously write their name and had a basic knowledge of the alphabet. We relied on reports from both parents and teachers as an accurate assessment of this ability. Consent forms were sent to the participating nursery or after school club and these were forwarded to parents. Those children who had obtained consent were then asked if they would like to participate in our study.

Age, gender, handedness and literacy were recorded for each child alongside spontaneous written productions of their names and letters of the alphabet. These were sampled using both the dominant and non-dominant hands. A perceptual measure of letter orientation discrimination was also included.

PROCEDURE

The children completed a set of writing tasks in which productions were recorded on a digitizing tablet. Children were seated at a desk with the touch screen tablet placed squarely in front of them. They were asked to write with both their dominant and non-dominant hands using their index finger, and stickers were used to help the children identify each hand. By requesting children to use their finger as opposed to a stylus, we were able eliminate any difficulty in holding a pen with the non-dominant hand. It was not always possible to keep the testing environment quiet or free from distraction and the location of testing also differed, but for the most part the testing environment was kept relatively constant.

Subjects were tested one at a time and consistent testing order was maintained across all participants. Testing of each subject took fifteen minutes to complete and was administered in the following order:

As a preliminary test to establish handedness, the children were asked to pick up the stylus and draw a circle on the tablet. The hand that they chose to draw with was considered to be their preferred, and therefore dominant, hand. The children were then asked to write their name spontaneously using the index finger of their dominant hand. This aspect of the procedure doubled as an initial literacy test. If the child could not spontaneously produce their name, they were excluded from the study. Children were then asked to write their name using their non-dominant hand. As a third preliminary measure, we administered a simple literacy test which comprised all 15 asymmetrical capital letters of the alphabet (B, C, D, E, F, G, J, K, L, N, P, Q, R, S, Z) printed on white A5 cards in 250 point size Times New Roman black font. The asymmetrical capital letters were presented to the child one at a time in alphabetical order and the subjects were asked to name each letter. For the younger children, making the noise of the letter was sufficient.

We carried out two experimental tasks, the first being the bimanual motor task, and the second the directional discrimination task. To investigate whether mirror writing could be attributed to motor factors, subjects were asked to write the asymmetrical letters that they could name, all initially with their dominant hand and then all with their non-dominant hand. They were asked to write both the upper and lower case of the letter if possible. Finally, to test the prominence of perceptual factors in mirror writing, subjects completed a perceptual task in which letters were presented in normal and mirrored form and the subjects had to indicate which orientation of each letter was correct. The same 15 asymmetrical letters were presented on the tablet using a specialized computer program. The letters were printed in black uppercase Arial font against a white background. The experimenters selected the asymmetrical capital letters that were known by the child, of a possible 15, and these were displayed in a randomized order, one at a time in the center of the screen. For the children who recognized 10 or more, each letter was presented in both a normal and mirrored orientation. For children who recognized less than 10, the letters were reused until a 10-letter list had been completed. There were therefore between 20 and 30 trials per child. Participant responses were recorded by the experimenters by pressing a button corresponding to either correct or incorrect direction.

Data Recording All written productions were recorded as bitmap images and saved to the digitizing tablet. The perceptual test responses were saved as text documents. A separate paper record of all production errors as well as perceptual errors was kept for all participants as a backup in case of a computer fault.

Data Analysis The data for each child was coded as follows: every mirror reversal of a letter was denoted by '1' for error; all correctly written letters were denoted by '0' (for both name and writing task). Similarly, every error made on the perceptual task (if a mirrored letter was said to be correct or vice versa) was denoted by '1', and again '0' represented a correct response. The literacy score was calculated as a proportion of the letters known by the child (of a possible 15). The proportion of mirror writing per child was calculated from the number of mirror reversals with respect to the number of letters written by that child. Perceptual error was scored in the same way.

RESULTS

Preliminary Analysis In order to assess the relationship between age, mirror writing and perceptual errors, the children were split into a mirror writing group and a non-mirror writing group. Those who did not produce any mirror writing in their written productions were excluded from the analysis (n = 15). The final analysis for mirror writing among children was carried out on 36 participants. Total proportionate mirror writing was calculated for each child, along with the proportion of mirror writing carried out with the dominant hand, non-dominant hand, lowercase letters and uppercase letters. 5 children did not complete the perceptual task. The proportion of perceptual errors was calculated for each of the remaining children (n = 46).

We conducted a similar analysis with respect to letters. For this, the proportion of mirror writing with the dominant hand and non-dominant hand was calculated for each letter. A motor score was derived from the average number of times that a letter was mirror written with both hands. A perceptual error score for each forward and backward facing letter was also calculated, and the total perceptual error score was derived from the average number of times each letter was confused. We assessed the relationship between these scores and included the direction of the letters as a grouping variable.

As there were only 7 instances of mirror written names, this variable was not considered pertinent to our investigation and was therefore excluded from the analyses.

The majority of children performed at ceiling level in the literacy test; therefore, this was not considered to be a practical assessment of language development and age was used instead for comparative purposes in tracking the progression of mirror writing and perceptual confusion.

Normality plots and Shapiro-Wilk tests indicated that all data was significantly non-normal, positively skewed and leptokurtic. This non-normality could not be rectified by an arcsine transformation and therefore non-parametric methods were used to analyze the data.

Statistical Analysis

Descriptive Statistics. The mean mirror writing per child gradually decreased with age (Table 1). A significant proportion of 4 year olds' writing was mirrored (24.4%); however, this was representative of a very small sample (n = 5). Less than 2% of writing was mirrored in children aged 8 years and older (n = 9). The proportion of perceptual errors for different age groups corresponded to the values of mirror writing, 4 year olds were inclined to make more perceptual errors than other age groups (38%) and this error rate gradually decreased with age, reducing to only 1.4% for ages 8 and over.

Research Question 1: Is there a relationship between age and mirror writing? Age was found to correlate significantly and negatively with both perceptual error [r_s = -.69, *df* = 46, *p* < .001] and mirror writing [r_s = -.47, *df* = 51, *p* < .001].
There was no significant effect of gender on either proportion of mirror writing (*U* = 308.5; exact *p* = .4) or proportion of perceptual errors (*U* = 261.5; exact *p* = .494). Girls had an average rank of 26.59 for mirror writing and 23.45 for perceptual error. Boys had an average rank of 25.52 and 23.54 for mirror writing and perceptual error respectively.

TABLE 1. Average rate of mirror writing (MW) and perceptual error (PE) per child, grouped by age

Age (Yrs)	Av. MW (n)	Av. PE (n)
4	24.4% (5)	38% (2)
5	17.7% (16)	23.2% (15)
6	11.4% (13)	13.7% (12)
7	9.2% (8)	5.5% (8)
> 8	1.7% (9)	1.4% (9)

Research Question 2: Is mirror writing attributable to perceptual confusion? In order to determine the association between mirror writing and perceptual errors, the proportion of mirror writing per child was compared to the average perceptual error for that child. Analysis by Spearman's rho showed a significantly positive correlation [r_s = .667, *df* = 46, *p* < .001], demonstrating that mirror errors increased with perceptual confusion. Given that the perceptual task consisted of only uppercase letters, we also correlated perceptual error with percentage of mirror written uppercase letters for each child to obtain a more accurate picture. These scores were again shown to be significantly positively related, r_s = .677, *p* < .001.

Research Question 3: Does directionality of letters predict mirror writing? Further to the evidence reported by Fischer (2011), we investigated whether left-facing letters of the alphabet were likely to be mirror reversed more often than right-facing letters. The letter J was mirrored a total of 37 times, while Z was mirror written 42 times. Both of these values were significantly above average (Table 2).

Similarly, the rate of mirror writing among letters was significantly positively correlated with the rate of perceptual confusion for that letter, r_s = .786, *df* = 15, *p* = .001.

When the perceptual confusion scores for letters were plotted against the proportion of mirror writing for that letter, a clear picture of Fischer's 'right writing rule' emerged (Figure 1).

A Mann Whitney U showed that mirroring and confusion of left (*M rank* = 14.5) and right (*M rank* = 7) facing letters was significantly different, *U* = .000, exact *p*(one tailed) = .01; r_g = .1, a 'large' effect by Cohen's (1988) classification. We can conclude from this that the left-facing letters J and Z were mirror written and perceptually confused considerably more often than the other 13 asymmetrical capital letters (Figure 2).

TABLE 2. Proportion of mirror writing instances for each letter

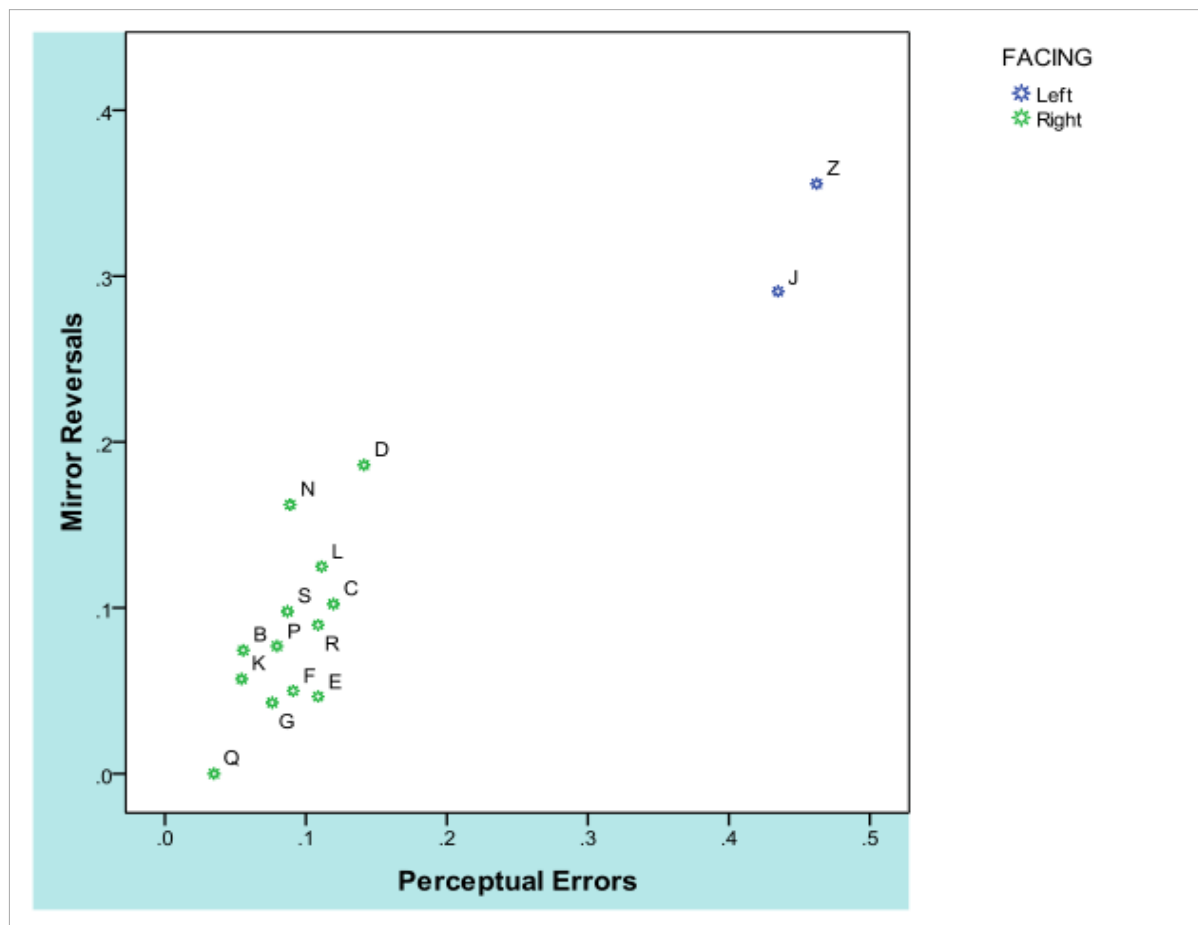| Letter	Average Instances of MW (%)
B	8
C	10
D	19
E	5
F	5
G	4
J	**29**
K	6
L	13
N	16
P	8
Q	0
R	9
S	10
Z	**36**

FIGURE 1. Mirror writing compared to perceptual errors for each letter, grouped by direction.
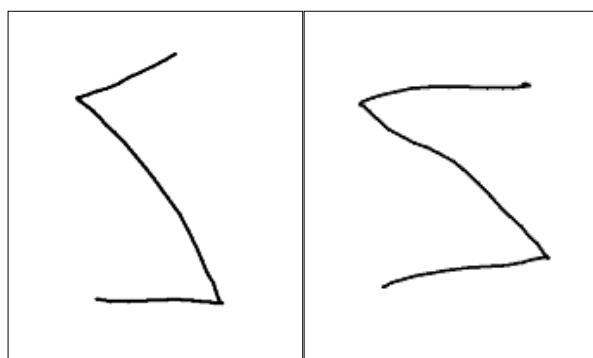

FIGURE 2. An example of perceptually driven mirror writing of Z by a child aged 81 months.

Research Question 4: Can mirror writing also be considered a motor phenomenon? In addition to the role of perception in mirror writing, we wanted to investigate if there are also motor processes involved.   The mean number of children in whom dominant mirror writing occurred more frequently than non-dominant mirror writing (M = 15) was higher than those in whom non-dominant mirror writing occurred more than dominant mirror writing (M = 11.25). However, this difference did not reach statistical significance so it was not possible to conclude a motor aspect to mirror writing from this analysis alone, (p = .402). The sums of ranks were 120 and 180 for negative and positive ranks respectively, therefore W = 120.

With respect to dominant and non-dominant hand mirror writing, there was a lower negative correlation between age and non-dominant hand mirror writing [r_s = -.405, df = 51, p < .01], compared to dominant hand mirror writing [r_s = -.519, df = 51, p < .001]. This indicates that non-dominant hand mirror writing may persist longer than dominant hand mirror writing as children grow older.

This non-dominant bias may be illustrative of a motor aspect of mirror writing.   To probe this theory, the age range was split into two groups along the median age (75.5 months among mirror writers).   Mirror writing in the younger group (n = 18) was shown to be biased towards the dominant hand, while in the older group (n = 18), it was biased towards the non-dominant hand (Figure 3).   Our index of bias was calculated by subtracting the mirror writing scores of the dominant hand from those of the non-dominant hand. Therefore, a positive value indicates a non-dominant bias and a negative value indicates a dominant hand bias. To test the statistical significance of this pattern, the two groups were compared against their hand bias for mirror writing.   The data was not normally distributed and was therefore analyzed by means of a Mann Whitney U test.   The average rank of the older age group (22.14) was greater than the average rank of the younger age group (14.86). This difference was shown to be significant, U = 97; exact p (one-tailed) = .017; r_g = .4, a "medium" effect (Cohen, 1988).   This shows that mirror writing in older children was biased to their non-dominant hand and is indicative of a gradual transference from perceptual to motor processes in mirror writing as children grow older.
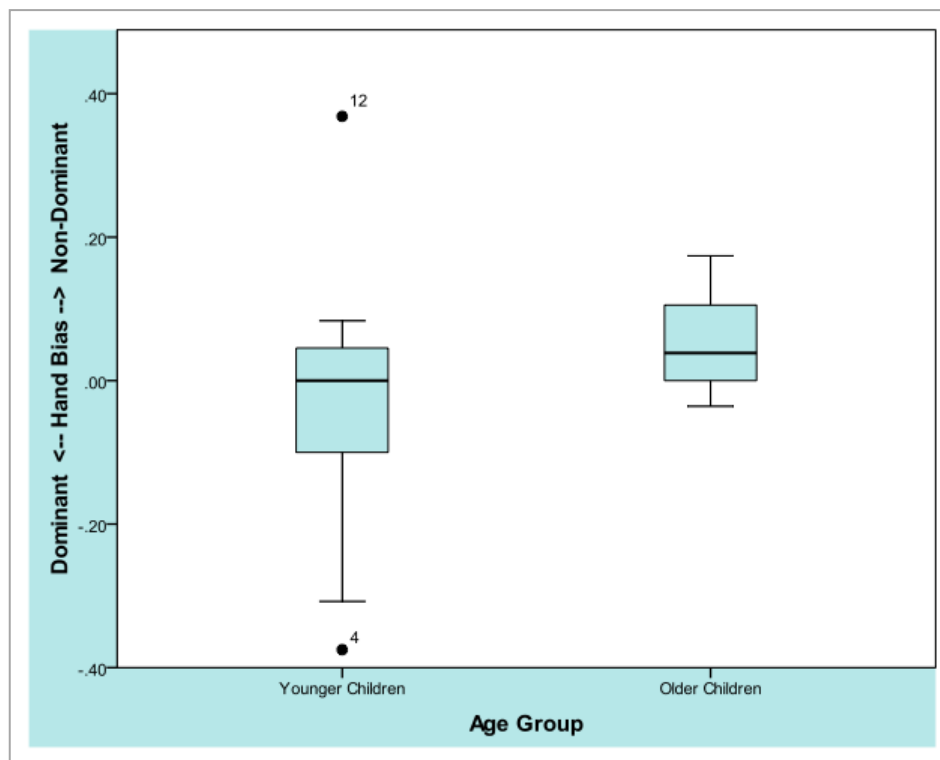

FIGURE 3. Hand bias for younger and older age groups.

DISCUSSION

Deciphering the behavioral basis of mirror writing has posed a significant challenge for researchers for over a century. Given the ongoing debate over perceptual and motor accounts of mirror writing, and with considerable evidence supporting each side of the argument, we posited a role for both a perceptual and motor foundation of mirror writing in children.

Our study provides compelling evidence for a perceptual basis to mirror writing. Perceptual errors were shown to significantly increase with the rate of mirror writing ($p < .001$). This finding is at odds with studies that have posited a primarily motor basis to this phenomenon in children (Della Sala & Cubelli, 2007). Della Sala and Cubelli found no relationship between perceptual error and mirror writing; however their measure of perception was an odd-one-out picture task. Our task of letter discrimination may be a better predictor of perceptual confusions with written stimuli and may explain this disparity in the results. Our findings are supportive of previous research with brain-damaged adults (Gottfried et al., 2003; Heilman et al., 1980) and control samples (Yang, 1997; Tankle & Heilman, 1983). However, despite the compelling evidence for a perceptual foundation to mirror writing, it only represents half of the story.

We found that perceptual errors and mirror writing both significantly decrease with age ($p < .001$). This is in support of Cornell's (1985) findings. A distinct parallel can be drawn across Cornell's findings and our own, with respect to the prevalence of mirror writing among different age groups. Cornell found that 82% of 5 year olds mirror wrote, while we found an 81% prevalence of mirror writing among 4 and 5 year olds. The high prevalence of mirror writing in recent studies such as this is at odds with older studies which reported mirror writing as extremely rare among children (Orton, 1928). These reports of low incidence rates led to the perception of mirror writing as an abnormality and a developmental dysfunction. In another similarity to this study, Cornell also found no effect of gender on mirror writing. Despite the high occurrence of mirror writing in younger children, Cornell found that only 13% of 7 year olds mirror wrote, while 76% of our 6 and 7 year olds produced mirror reversals. In fact, the prevalence of mirror writing in our study only dropped as low as 33% for children aged between 8 and 10 years. This discrepancy may be best explained by the different methods used. In Cornell's study, a spatial constraint was used to induce mirror writing in children. The persistence of mirror writing in our experiment may be attributable to the fact that we used a motor technique to elicit mirror writing as opposed to a perceptual one.

We also established that perceptual errors in mirror writing gradually dissipate with age and are replaced by a motor mechanism. This transition from perceptually-dominated to motor-driven mirror writing is illustrated as occurring between the ages of 6 and 7 years. The basis of mirror writing before this age differs significantly with the motor patterns of older children ($p = .017$). This development is at odds with Della Sala and Cubelli's (2007) suggestion of directional apraxia. In our study, experience with language in young children seems to initially shift the causes of mirror writing from perceptual to motor factors, rather than abetting the acquisition of a correct motor direction. Although the theory of directional apraxia may apply to dominant hand mirror writing, non-dominant hand mirror writing almost certainly seems to be the execution of a learned motor program in mirrored form. This finding lends substantial support to the motor hypothesis and corroborates evidence from pathological cases of mirror writing (Balfour et al., 2009; Buxbaum et al., 1993). This motor mechanism is not directly observable in adults as sufficient cognitive control strategies are assumed to be in place to override such a phenomenon.

This study also substantiated claims made by Fischer (2010, 2011) in that the direction of letters is a major factor in relation to both mirror writing and perceptual confusion. We found that the leftward facing letters J and Z were mirrored and confused significantly more often than rightward facing letters ($p = .01$).

Two children stand out as good illustrations of each of these processes, one exhibiting a near perfect perceptual pattern (aged 78 months) and the other a near perfect motor pattern (aged 64 months). In isolation, the ages of these children conflict with our finding that mirror writing is increasingly influenced by motor factors as children grow older, however, these children attended different schools and as such, this discrepancy is probably attributable to individual differences in their experience with writing.

In the case of the younger child, the majority of capital letters were written in the correct direction with the dominant hand, with the exception of J, R and S. A large proportion of the letters that were written correctly with the dominant hand were then mirrored with the non-dominant hand, with the exception of N, P, Q and Z, which were written correctly with both hands, and R and S which were written mirrored with both hands. An interesting observation is that J was mirrored with the dominant hand and then written correctly with the non-dominant hand (Figure 4). This indicates that perceptual and motor influences are operating simultaneously. It seems that J may have been perceptually confused to begin with but the mirrored motor program unintentionally rectified this confusion. The perceptual error score for this child was 63%
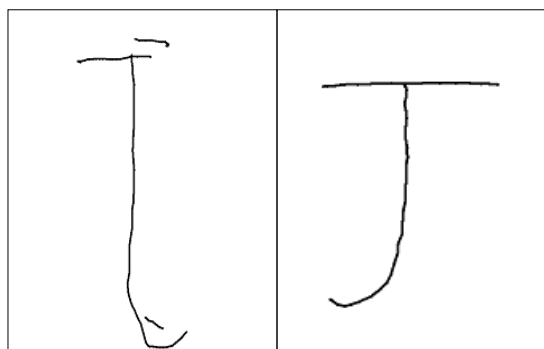
FIGURE 4. Child's mirror written Js with dominant hand (left) and non-dominant hand (right).

which may show that the child was on the verge of overcoming perceptual confusion and was still vulnerable to motor reversals for most letters. It is likely that perceptual factors are still impacting the letter J as it is a leftward facing letter and therefore has a higher perceptual confusion rate. Another interesting observation is that Z was written correctly with both hands, demonstrating that perceptual confusions were diminishing.

The second interesting case exhibited a predominantly perceptual pattern. Although this child was 14 months older, she seemed to have persistent perceptual confusion in writing, as all letters with the exception of B, C, J, Q and Z were mirrored consistently with both hands (Figure 5).
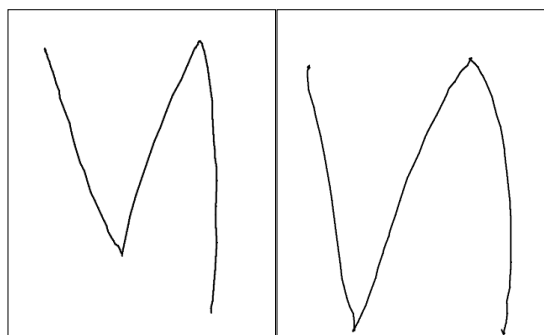

FIGURE 5. Example of mirror written Ns with dominant (left) and non-dominant hand (right).

This is interesting as her correct writing of J and Z violate the frequency with which these letters are usually mirrored relative to right facing letters. A noteworthy factor of this case was that the child had a relatively low perceptual error score of 33%. This may highlight a possible discrepancy in the development of perceptual mechanisms governing mirror writing and mirror reading. It is possible that mirror discrimination abilities may mature slightly faster than writing abilities in certain children. This theory would only be testable by means of a longitudinal study.

Previous studies in this area have generally lacked a cohesive approach to examining the fundamentals of mirror writing. The literature has largely focused on satisfying one side of the perceptual/motor debate or the other, which has limited the more exploratory stance of considering both mechanisms simultaneously. Despite the spontaneous occurrence of mirror writing in children, the majority of research in this area has concerned pathological cases of mirror writing or intentional mirror writing in adults. These studies are highly repetitive with regards to both the methods and participant sample used.

The study of mirror writing over the past few decades has been restricted by previous assumptions and has only recently received renewed attention. New perspectives on the phenomenon in children have been introduced in recent times using brain imaging technology (Dehaene et al., 2010) and novel theories (Fischer & Tazouti, 2012; Della Sala & Cubelli, 2007). Concerning the theory of directional apraxia, the conclusions reached from these studies were inferred from inconclusive methods. As discussed above, the perceptual task devised by Della Sala and Cubelli (2007) perhaps failed to accurately assess perception of mirrored letters, and the use of inappropriate stimuli led to perceptual theories of mirror writing being discounted. This study appears to lack a comprehensive approach to assessing motor influences of the non-dominant hand on mirror writing in children.

Our approach adopted a more exploratory technique and addressed some of the gaps in the literature. Up to this, no study had explored the impact of non-dominant hand writing in children and this process was made easier by allowing children to write with their fingers on the tablet. This novel approach meant that children did not have to concentrate on holding a pen with an unfamiliar hand and may have resulted in a more natural writing output with the non-dominant hand. Assuming that this natural script is mirrored, this technique is therefore the best measure of the motor hypothesis. This study has also contributed significantly to the literature by the use of a letter perceptual discrimination task. The use of both normally oriented and mirrored letters in this task allowed us to thoroughly gauge the child's reactions to both forms of letters and it was noted that children showed equal levels of confusion with both correctly oriented and mirrored letters.

These results contribute a great deal to the current debate surrounding mirror writing. Despite this, our study was limited in that our sample was relatively small and the participants recruited were from similar backgrounds and education systems. It would be valuable to test these findings not only with bigger samples, but also in different cultures or with left facing

languages, to determine the universality of our findings. Our approach would also be furthered by replication with a larger sample of left handers (n = 5 in the present study) to fully assess the implications of the motor hypothesis. This may succeed in corroborating the evidence stated here, but such a study may also undermine our findings if the motor hypothesis is demonstrated to be specific to the left hand as opposed to the non-dominant hand.

More specific limitations of our study stem from aspects of our design that would benefit from modification. A more thorough measure of literacy is essential for assessing different stages of language acquisition in children and would provide a more suitable measure for tracking the development of mirror writing, rather than age. Our perceptual task ought to include lower case letters as these were written in conjunction with upper case letters during the writing tasks. There is also a need for consistency with the fonts used for both the literacy and perceptual tasks. The discrepancy between the *serif* Times New Roman and the *sans-serif* Arial fonts sometimes created confusion, particularly with upper case mirrored J. In Arial font, this letter is ambiguous as it also resembles lower case L and during the perceptual task, we had to reiterate to the children that this letter was a J. Our findings added credence to Fischer's "right writing rule" with respect to letters. However, it may also be beneficial to include digits in future mirror writing studies to gain a more comprehensive view of this mechanism (Fischer, 2011).

There is also room for improvement with our assessment of handedness. It was noted that some children seemed equally comfortable using both hands to write and would sometimes attempt to switch between their preferred and non-preferred hands during testing. One child, who preferentially wrote with his right hand initially, reported that it was "more comfortable" to write with his left hand during testing. We also noted that difficulties with the use of the tablet as on some occasions the children had to make several attempts before the tablet registered their writing. As a result, when writing with the non-dominant hand, some of the children's original productions were traced in a mirrored direction. When writing was disrupted, leading to them having to concentrate harder, some children corrected their direction and wrote the letter normally. We only included the final production of the letter; thus, samples of mirror writing were lost due to this difficulty.

It is evident that mirror writing may be more common and may persist longer in older groups of children than previously thought. As such, future studies may wish to track this phenomenon in children as they progress into adolescence. Replica studies by longitudinal analysis are needed to determine the robustness of our findings. This method would be paramount in tracing the progression of mirror writing throughout childhood. Age and literacy cannot fully account for the development of mirror writing as children mature at different rates as language and writing skills are impacted by a large number of contributing factors, including influences from the home environment, schooling, parents and siblings. Continued convergence of neuropsychological evidence, experimental studies with children and development of new methods of analysis, such as brain imaging techniques as well as longitudinal research, is necessary to gain a comprehensive understanding of the impact of these factors. Only then will we be able to fully grasp the basis of this intriguing phenomenon.

References

Angelillo, V. G., De Lucia, N., Trojano, L., & Grossi, D. (2010). Persistent left unilateral mirror writing: A neuropsychological case study. *Brain and Language, 114*, 157-163.

Balfour, S., Borthwick, S., Cubelli, R., & Della Sala, S. (2007). Mirror writing and reversing single letters in stroke patients and normal elderly. *Journal of Neurology, 254*, 436-441.

Brown, J. S., Knauft, E. B., & Rosenbaum, G. (1948). The accuracy of positioning movements as a function of their direction and extent. *American Journal of Psychology, 61*, 167-182.

Buxbaum, L. J., Coslett, H. B., Schall, R. R., McNally, B., & Goldberg, G. (1993). Hemispatial factors in mirror writing. *Neuropsychologia, 12*, 1417-1421.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed.).* Hillsdale, NJ: Lawrence Earlbaum Associates.

Cornell, J. M. (1985). Spontaneous mirror-writing in children. *Canadian Journal of Psychology, 39*, 174-179.

Critchley, M. (1928). *Mirror Writing.* London: Kegan Paul, Trench, Trubner & Co. Ltd.

Cubelli, R., & Della Sala, S. (2009). Mirror writing in pre-school children: a pilot study. *Cognitive Processing, 10*, 101-104.

Dehaene, S., Nakamura, K., Jobert, A., Kuroki, C., Owaga, S., & Cohen, L. (2010). Why do children make mirror errors in reading? Neural correlates of mirror invariance in the visual word form area. *Neuroimage, 49*, 1837-1848.

Della Sala, S., & Cubelli, R. (2007). 'Directional Apraxia': A unitary account of mirror writing following brain injury or as found in normal young children. *Journal of Neuropsychology, 1*, 3 - 26.

Fischer, J. P. (2010). Capital letters mirror writing in children: a new hypothesis to unlock the mystery. *Enfance, 62*, 371-386.

Fischer, J. P. (2011). Mirror writing of digits and (capital) letters in the typically developing child. *Cortex, 47*, 759-762.

Fischer, J. P., & Tazouti, Y. (2012). Unraveling the mystery of mirror writing in typically developing children. *Journal of Educational Psychology, 104*, 193-205.

Gottfried, J. A., Sancar, F., & Chatterjee, A. (2003). Acquired mirror writing and reading: Evidence for reflected graphemic representations. *Neuropsychologia, 41*, 96-107.

Heilman, K. M., Howell, G., Valenstein, E., & Rothi, L. (1980). Mirror-reading and writing in association with right-left spatial disorientation. *Journal of Neurology, Neurosurgery and Psychiatry, 43*, 774-780.

Noble, J. (1968). Paradoxical interocular transfer of mirror-image discrimination in the optic chiasm sectioned monkey. *Brain Research, 10*, 127-151.

Orton, S. T. (1928). Specific reading disability - strephosymbolia. *Journal of the American Medical Association, 90*, 1095-1099.

Paradowski, W., & Ginzburg, M. (1971). Mirror writing and hemiplegia. *Perceptual Motor Skills, 32*, 617-618.

Pedago, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011). Breaking the symmetry: Mirror discrimination for single letters but not for pictures in the Visual Word Form Area. *Neuroimage, 55*, 742-749.

Rodríguez, R. (1991). Hand motor patterns after the correction of left-nondominant-hand mirror writing. *Neuropsychologia, 29*, 1191-1203.

Rodríguez, R., Aguilar, M., & González, G. (1989). Left non-dominant hand mirror writing. *Brain and Language, 37*, 122-144.

Schott, G. D. (1999). Allen's self observations, Lewis Carroll's "looking-glass" letters, and Leonardo da Vinci's maps. *Lancet, 354*, 2158-2161.

Schott, G. D. (2007). Mirror writing: neurological reflections on an unusual phenomenon. *Journal of Neurology, Neurosurgery and Psychiatry, 78*, 5-13.

Tankle, R. S., & Heilman, K. M. (1983). Mirror writing in right-handers and left-handers. *Brain and Language, 19*, 115-123.

Tucha, O., Aschenbrenner, S., & Lange, K. W. (2000). Mirror writing and handedness. *Brain and Language, 73*, 432-441.

Wang, X.-d. (1992). Mirror writing of Chinese characters in children and neurological patients. *Chinese Medical Journal, 105*, 306-311.

Yang, M. -J. (1997). Mirror writing in right-handers and in left-handers: a study using Chinese characters. *Neuropsychologia, 35*, 1491-1498.

Openness to Experience and Health: A Review of the Literature

Lameese Eldesouky
University of California, Berkeley

The Five Factor Model, one of the most commonly used models for assessing personality, consists of five main universal traits: extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience (McCrae & John, 1992). Recently, its application to fields investigating the role of personality on physical and mental health has been extensive, with most traits producing consistent results across studies. The fifth of the model's main traits, however, openness to experience, has produced inconsistent results. In this review, we attempt to understand why these findings have been mixed by analyzing various facets of this trait in depth. We evaluate the six facets of openness to experience: actions, ideas, values, aesthetics, fantasy, and feelings individually to better understand the implications of openness to experience on physical and mental health.

Personality can be defined as the consistent behavioral, attitudinal, and motivational patterns that differ across individuals (McCrae & John, 1992). To describe this enduring set of patterns, theorists have long sought ways to identify and categorize the basic elements of personality. One of the primary methods to identify and categorize these basic elements has been factor analysis, where thousands of personality traits are listed and then eliminated as similar traits are grouped together (McCrae & John, 1992). As the field of personality psychology has progressed, though, theorists have continually disagreed on what the basic dimensions are and how many there should be. And while there still remains wide disagreement across theorists, one recent framework has been particularly used and accepted across psychology and related fields. This model is known as the Five Factor Model.

Evidence for the Five Factor Model began to make its way into personality psychology in the early 1980s. Psychologists Paul Costa and Jeff McCrae had finished reviewing multiple personality scales and developed an integrative personality scale known as the NEO Personality Inventory Revised (NEO PI-R), which measured five particular traits: neuroticism, extraversion, openness to experience, conscientiousness, and agreeableness (John, 2008). The appeal of their work came largely from its ability to take into account factors that were similar across different personality scales (John, 2008). For instance, extraversion, the trait describing one's propensity towards sociability and positive emotions, and neuroticism, the trait describing emotional stability, can be found in earlier personality scales such as Allport's Trait Theory, Cattell's Sixteen Personality Factors, and Eysenck's Big Three (McCrae & John, 1992). Thus, while personality psychologists frequently disagreed over what the basic dimensions were and their number, there was some agreement on certain dimensions. Following the Five Factor Model research done by Costa and McCrae and other researchers, additional personality scales measuring these five particular traits were later developed, including Lew Goldberg's International Personality Item Pool, Gerad Saucier's Big Five mini-markers, and Oliver John's Big Five Inventory (BFI) (John, 2008).

While researchers differ in which questionnaire they may prefer to use, many of the questionnaires have proven to be reliable and only have a few discrepancies in questions and length.

The Five Factor Model, also labeled the Big Five, does not refer to a particular questionnaire, but rather a general personality framework based off of five main universal traits: extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience (John, 2008). Extraversion, as briefly mentioned above, is one of the most studied personality traits, probably due to the large role it plays in our interpersonal lives (John, 2008). It is a trait characterized by high energy, enthusiasm, and assertiveness. A behavioral example is of extraversion is an individual who makes the effort to approach strangers and begin conversation. An individual low on the trait of extraversion would be less likely to perform such an action. Less extraverted individuals often prefer to keep to themselves and engage in solitary activities. While low extraversion does not necessarily imply anti-sociability, less extraverted individuals generally have smaller social circles than their more extraverted counterparts.

Neuroticism, also briefly mentioned above, deals with emotional stability. In fact, some personality theorists such as Cattell had referred to it directly as emotional stability (McCrae & John, 1992). Like all of the Big Five traits, neuroticism is also on a spectrum. Those who are high on the trait are more prone to negative feelings such as anxiety, tension, and depression, while those who are low on the trait are calmer and more even-tempered (John, 2008). A behavioral instance of neuroticism is an individual who might get overly anxious when there is little to be anxious about and ruminate continuously about negative events that have already occurred. A less neurotic individual would relax about most situations and when faced with a difficult situation, try to see the good in it.

Conscientiousness, the third trait, describes order, discipline, and impulse control (John, 2008). It is most noted for its ability to influence the organization and direction of behavior (McCrae & John, 1992). High conscientiousness can be seen in people who are punctual, have a strong work ethic, do not get easily distracted, and take good care of their health by eating right and taking their medication (John, 2008). They are people who are focused, prepared, and

almost always have a plan. Individuals low on conscientiousness are unorganized, easily distracted, and often not as disciplined. They prefer spontaneity and are not as concerned about controlling aspects of their environment or even their own behavior.

Agreeableness is the fourth of the Big Five traits. It is similar to extraversion in that it is deeply related to interpersonal relationships. However, it is more concerned with the way one treats and deals with others, as opposed to whether or not one actively seeks relationships out. For instance, one who is highly agreeable is empathetic towards others. They are affectionate, trustworthy, and altruistic (John, 2008). Due to their kind nature, they are able to work better in groups and often have relationships with little conflict (John, 2008). Less agreeable individuals may be aggressive, rude, and look at others with contempt. Unlike highly agreeable individuals, they are not very considerate of people's emotions and might be far more concerned about themselves and their own desires. Their lack of consideration for others might result in unsatisfying relationships and numerous interpersonal problems.

The last of the Big Five traits is openness to experience, which is described by novelty-seeking, intellectual curiosity, a vivid imagination, awareness of inner emotional states, and deep appreciation for the arts (John, 2008). Individuals high on this trait seek out activities that bring meaning to their lives and allow them to think about things in a different way. They may prefer to change their routine to make things more interesting or they might seek out an opportunity to do something they have never done before. Individuals low on this trait prefer sameness and predictability and are often unoriginal and closed-minded. Overall, they are less likely to seek out activities that may enrich their mental and experiential lives, primarily because they are not interested in doing so.

Application of the Five Factor Model to Health

The comprehensiveness of the Big Five traits has enabled researchers to better use personality traits as predictors for certain types of behaviors or outcomes, and thus has been applied to numerous fields including education, industrial and organizational psychology, developmental psychology, and forensics

(McCrae & John, 1992). More recently however, the Big Five have been used in the field of health, with a large focus being placed on understanding how personality factors can affect mental and physical health. Extraversion, conscientiousness, and agreeableness for instance, have been linked to positive outcomes such as greater immune response to infection (Capitanio, Abel, Mendoza, Blozis, McChesney, Cole, & Mason, 2008; Sloan, Capitanio, Tarara, & Cole, 2008), lower probability of developing a cold (Cohen, Doyle, Turner, Alper, & Skoner, 2003), positive health perception (Jerram & Coleman, 1999), longevity (Bermudez, 1999), and positive adjustment to illness or disability (Boyce & Wood, 2011). Neuroticism on the other hand, has been linked to higher disease risk and other negative outcomes such as decreased levels of oxytocin and neuropeptide Y (Schaller & Murray, 2011).

When it comes to understanding the role of openness to experience in health, findings have been mixed, with some studies finding that high openness to experience is beneficial and others finding that it can contribute to negative outcomes. Perhaps the challenge in making sense of the relationship between health and openness to experience arises from the controversial nature of the trait itself. Most personality psychologists endorse some form of openness to experience, but throughout the history of personality psychology the trait has continued to take on different forms (John, 2008). Some have referred to it as an indicator of intellect, while others have referred to it as an indicator of creativity.

Applying a Facet Approach. In the longest and most detailed of the Big Five questionnaires, the NEO-PI-R, Costa and McCrae have made individual distinctions for each trait within their survey by assigning every trait six *facets* to better understand the trait (McCrae & John, 1992). Facets can be considered as more detailed dimensions of an individual personality trait. For instance, one may be extraverted in the sense that they are very assertive, while one may be extraverted in the sense that they are gregarious. It is not necessary for an individual who scores high on a Big Five trait to embody all of its characteristics. Thus, in this example of extraversion, Costa and McCrae divide aspects of the trait into facets, two of which are gregariousness and assertiveness.

Facets of the Big Five traits are meant to be inter-related and yet despite being grouped under the same trait, the facets of openness to experience in particular have been argued to almost be individual personality traits themselves. The six facets of openness to experience are openness to actions, ideas, values, aesthetics, fantasy, and feelings (Coan, 1972). Openness to actions describes the drive to seek out new activities and attempt new things; openness to ideas is the desire to be intellectually curious and think about things in new and interesting ways; openness to values is the willingness to re-examine one's traditional values, be it political, cultural, or religious; openness to aesthetics describes the tendency for one to appreciate the arts; openness to fantasy is the proclivity towards deep imagination and fantasy; and the last facet, openness to feelings, describes how in tune one is with their emotional states.

While individuals who score high on openness typically score high on most of the facets of openness, the distinction of each facet makes it possible for someone to be very open in one area, but very closed in another (Coan, 1972). For instance, consider the difference between the person who seeks out exhilarating activities such as bungee-jumping and the person who prefers calmer activities such as attending art exhibits. Even if they are both high on openness, they can vary widely in the types of experiences to which they are open (Coan, 1972). Thus, it is not accurate to say that such individuals are open in the same way, even if their overall openness scores turn out to be identical.

Since openness to experience is such a complex trait, it is reasonable to take into account each facet level when approaching research on the influence of the trait on health. Many studies that have examined the influence of openness on health have looked at the trait broadly instead of at the facet level. It may be that failure to take into consideration scores on the individual facets makes results challenging to interpret. In this review, we attempt to understand what the physical and mental health implications of openness are by deeply examining the individual facets. We first group the six facets into two main categories: facets that have a strong behavioral and cognitive component, openness to actions, ideals, and values, and facets that have a strong emotional component, openness to aesthetics, fantasy, and feelings. The reason for this grouping is because facets vary in their main focus and it is simpler to discuss facets that are similar in their

focus than facets that are different. For instance, the facet of actions has a strong behavioral component because it is geared towards seeking out physical activities and is less concerned with emotion, whereas the facet of feelings has a strong emotional component because it describes how much one values their emotional experience and is less concerned with physical pursuit of certain activities. When discussing each facet we will provide a description of the characteristics that the facet entails then discuss the relevant biological correlates that have been found thus far. Next, we summarize findings on links between each facet and health, with consideration for both physical and mental disorders. Finally, we review current controversies and theoretical difficulties, then propose some future directions for research.

The Facets of Openness to Experience

Openness to actions Openness to actions is characterized by involvement in varied experiences and enjoyment of novelty. It has been considered by McCrae and Costa (1997) to be a facet of motivation to engage in novelty and complexity. While it has been found to have a strong relationship with the trait of extraversion, McCrae and Costa (1997) have argued that it reflects a pure behavioral exploratory tendency, while openness to actions takes into account a cognitive element as well. To better understand the behavioral component of openness to actions, many researchers have also examined sensation seeking (Aluja, Garcia, & Garcia, 2002). This is because extraversion and openness to actions are often positively correlated with sensation seeking. Sensation seeking has been defined by Zuckerman (1979) as "the need for varied, novel, and complex sensations and experiences and the willingness to take physical and social risks for the sake of such experience." It has been linked to the seeking of high stimulating activities such as exotic meals, various sports, as well as illegal activities. Such activities include the willingness to take various risks to experience them and thus, frequently carry with them the trait of impulsivity (Gerra, Avanzini, Zaimovic, Satori, Boochi, Timpano, Zambelli, Delsignore, Gardini, Talarico, & Brambilla, 1999).

Biological correlates of openness to actions. In investigating the biological correlates of openness to actions, most studies have not directly assessed openness. Instead they have investigated the biological correlates of extraversion or sensation seeking. Both sensation seeking and extraversion have been linked to high levels of norepinephrine, low levels of monomamine oxidase (MAO), and variations in dopamine receptors (Cloninger, 2000). In investigating the relationship between norepinephrine and sensation seeking, it has been found that higher levels of norepinephrine may not be the cause, but rather a consequence of sensation seeking.

Norepinephrine is a catchechloamine that plays multiple roles, most importantly in the stress response. While norepinephrine levels increase with higher levels of cortisol, it has been hypothesized that only increased levels of norepinephrine, and not cortisol, are directly correlated with sensation seeking (Gerra et al., 1999). However, other research findings have presented a similar hypothesis, but under the belief that high sensation seeking is linked to lower and not higher levels of norepinephrine (Zuckerman, 1995). It has been proposed that high sensation seekers are chronically under-aroused and thus through the seeking of stimulating activities, are able to raise their levels of norepinephrine (Zuckerman, 1995). Likewise, other researchers have made a similar proposal with dopamine, proposing that individuals high on sensation seeking have high scores because they have an increased sensitivity of postsynaptic dopamine receptors and that they require higher densities of dopamine to overcome the sensitivity (Gerra et al., 1999).

In examining MAO, researchers have found that there is a negative correlation with the trait of sensation seeking (Zuckerman, 1995). MAO is an enzyme that regulates monoamine levels by breaking down various neurotransmitters. The type B MAO has been particularly noted for its significance because it breaks down dopamine (Zuckerman, 1995). Since low levels of MAO are related to sensation seeking and extraversion, it is not surprising that individuals high on sensation seeking and extraversion have higher levels of circulating dopamine. Research on gender differences in MAO levels has found that men typically have lower levels of MAO. In addition, males that carry the 3-VNTR MAOA gene variant have higher levels of MAO and correspondingly, significantly lower scores in openness (Samochowiec, J., Syrek, Michal, Ryzewska-Wodecka, Samochowiec, A.,

Horodnicki, Zakrzewska, & Kucharska-Mazur, 2004).

Another enzyme partly responsible for metabolizing catecholamines that has been briefly investigated is the enzyme catechol-O-methyltransferase (COMT). Unlike the findings on MAO, high enzyme activity of COMT has been shown to be related to high levels of extraversion and sensation-seeking (Reuter & Hennig, 2003). Reuter and Hennig (2003) have vouched for the functional polymorphism of COMT, VAII58MET, as a candidate gene locus to examine further in determining biological correlates of extraversion and related characteristics and behaviors. In their experiment they found that the group with the higher extraversion scores had the VAL/VAL polymorphism (Reuter & Hennig, 2003).

Most work on high levels of dopamine has consistently shown that it is linked to characteristics such as impulsivity, excitability, and a desire to explore (Epstein, Novick, Umansky, Priel, Osher, Blaine, Bennett, Nemaov, Katz, & Belmaker, 1996; Panksepp, 1998; Reuter & Hennig, 2003). Lower levels have been affiliated with less flexibility, but a calmer temperament (Epstein et al., 1996). The belief is that dopamine regulates the motivational component of openness similar to how it regulates extraversion (Depue & Collins, 1999). Research on Parkinson's patience has corroborated this hypothesis (Kaasinen, Nurmi, Bergman, Eskola, Solin, Sonninen, & Rinne, 2001). Patients with Parkinson's disease experience degeneration of motor movement due to the death of dopamine-generating cells in the brain's substantia nigra. Kaasinen and colleagues (2001) compared un-medicated Parkinson's disease patients and controls to see if there was a difference in sensation-seeking levels. Patients with Parkinson's disease were found to have lower sensation seeking scores than controls, although this is identified as an effect of the disease, not a cause (Kaasinen et al., 2001). Similar relationships between sensation seeking and dopamine have also been found in patients with schizophrenia, who unlike Parkinson's disease patients have abnormally high levels of dopamine and thus, higher levels of sensation seeking, as well as high openness scores (DeYoung, Peterson, & Higgins, 2003).

Epstein and colleagues examined the dopamine D4 receptor (D4DR) exon III polymorphism and found that it was linked to sensation seeking, with individuals that carry the 7 repeat allele having significantly high sensation seeking scores (Epstein et al., 1996). Okuyama and colleagues have said that the effects of the 7 repeat allele is a result of differences in ligand affinity (Okuyama, Ishiguro, Nankai, Shibuya, Watanabe, & Arinami, 2000). They have also looked at DRDR, namely a polymorphism at -521C/T, and found that individuals with the T variant of the C-521T polymorphism have reduced transcriptional efficiency (Okuyama et al., 2000). Thus, subjects with a T/T genotype had the lowest sensation seeking scores, while subjects with a C/C genotype had the highest sensation seeking scores (Okuyama et al., 2000). In addition to DRD4, Ishiguro and colleagues have also looked at the dopamine D2 receptor gene (DRD2) and found that high sensation seeking is correlated with the A2 allele. They state that this is likely a result of increased dopamine D2 receptor binding (Okuyama et al., 2000).

DeYoung and colleagues have attempted to examine possible differences between extraversion/sensation seeking and openness by proposing that extraversion is linked to dopaminergic projections to the striatum and that openness is linked to dopaminergic projections to the prefrontal cortex (PFC) and the anterior cingulate cortex (ACC) (DeYoung et al., 2003). They have argued that openness is linked to the PFC because the PFC plays numerous roles in cognitive function, particularly working memory, which is necessary for manipulating information and carrying out important executive functions (DeYoung et al., 2003). In addition, increased dopaminergic activation in the PFC has been linked to an improvement of performance on tests of cognitive ability and flexibility. Despite their proposal, however, after administering cognitive tasks such as letter randomization, word fluency, recency judgments, and spatial and non-spatial conditional association tasks, DeYoung and colleagues have found that openness to actions is the least strongly related openness facet to cognitive variables (DeYoung et al., 2003).

Health and openness to actions. Research on the links between health and openness to actions has given mixed results, with some evidence demonstrating a positive association and other evidence demonstrating a negative association. Evidence demonstrating a positive association has supported the hypothesis that individuals high on openness to actions have better physical health because they have better

mental health. Researchers in support of this proposal have argued that this is because individuals higher on openness to actions are more likely to engage in behaviors that make them happy (Salovey, Rothman, Detweiler, & Steward, 2000). This is particularly significant for individuals who are diagnosed with a serious physical illness. Seeking activities that improve a patient's mood can allow them to have greater confidence in themselves and recovery of their illness (Salovey et al., 2000). For instance, when examining patients with Parkinson's disease, it has been found that lower scores on sensation seeking have been linked to higher depression scores than controls (Menza & Mark, 1994).

Dua (1990) has found that openness to actions predicts "emotional stability, lack of depression, and positive affect, both from thoughts and from real-life experiences". In a more recent study, Garcia and colleagues found that high sensation seeking was linked to less depression (Garcia, L.F., Aluja, Garcia, O., & Cuevas, 2005). Carrillo and colleagues (2001) have also come across similar results, with openness to actions negatively correlating with the Neuroticism factor and being predictive of depression. Higher scores on openness to actions were significantly less associated with both neuroticism and depression (Carrillo, Rojo, Sanchez-Bernardos, & Avia, 2001). Similar findings have also been found in relation to gender differences. Males, who naturally have higher levels of dopamine, typically score much higher than women on openness to actions and are less susceptible to depression (Carrilo et al., 2001). Other evidence on the relationship between depression and openness to actions has also pointed to polymorphisms in the serotonin transporter, which has been linked to depression in several studies (Samochowiec et al., 2004). Women who have the short variant of 5-HTT-linked polymorphic region (5-HTT-LPR) have been found to be more susceptible to depression and have lower scores of exploratory excitability (Samochowiec et al., 2004).

Oswald and colleagues (2006) have examined the association between openness to actions and cortisol responses as indicators of stress and anxiety. Participants underwent a laboratory psychological stress test and had their cortisol levels measured before and after (Oswald, Zandi, Nestadt, Potash, Kalydijan, & Wand, 2006). Individuals who were lower on openness to actions had higher cortisol levels and individuals who were higher on openness had lower cortisol levels (Oswald et al., 2006).

Another study by Schneider and colleagues (2011) came across the same findings. This may be further evidence for the hypothesis that higher sensation seeking scores are linked to high levels of norepinephrine, but not high levels of cortisol.

While there is evidence in support of the benefits of openness to actions, such as proactive seeking of diverse experiences and a decreased risk for depression, other evidence demonstrates that high openness to actions can also be detrimental to health. Booth-Kewley and Vickers, Jr. (1994) conducted an experiment on the associations between openness to experience and health behavior. Several individuals high on openness to actions reported greater substance risk taking. Booth-Kewley and Vickers, Jr. also measured other domains of personality, but found openness to be the only significant personality domain that predicted substance risk taking (Booth-Kewley & Vickers, Jr., 1994). One common model that attempts to link personality and disease holds that personality can lead to disease through the practice of unhealthy behaviors (Friedman & Booth-Kewley, 1987). Since openness to actions demonstrates a positive association with substance risk taking due to the desire to explore and bring about pleasure, it may present a negative link with healthy behavior (Jerram & Coleman, 1999; Salovey et al., 2000).

Openness to Ideas

While openness to actions expresses the motivation behind the behavioral dimension of openness to experience, openness to ideas focuses more on the motivation behind the cognitive component. Individuals who are high on openness to ideas exhibit greater flexibility in terms of "processing information and exploring the environment" (DeYoung et al., 2003). As Batey and Furnham (2006) state, extraversion and openness to actions "predict only the quantity, not the quality of ideas". Individuals high on openness to ideas are willing to try to new things and examine and reflect on their new experiences (McCrae & Costa, 1997). Even though an individual high on openness to actions is likely to be high on openness to ideas, openness to ideas does not require the behavioral component of openness to actions. High scores of openness to ideas may or may not reflect willingness to engage in novel behaviors, but always indicate a greater

engagement in activities that have the potential to increase knowledge (Wainwright, Wright, Luciano, Geffen & Martin, 2008). The willingness to engage in activities that increase any kind of knowledge has led to findings on how openness to ideas affects intelligence and creativity (or plasticity).

When first determining the characteristics associated with openness to experience, McCrae and Costa had decided that intellect was related to the trait, but that it was too broad (McCrae & Costa, 1997). Amongst the Big Five traits though, openness is the only trait positively correlated with the intelligence quotient (IQ), which is frequently used to assess general intelligence (DeYoung et al., 2003). Of the six facets, openness to ideas has appeared to capture intellect the most adequately, with individuals high on this facet demonstrating greater efficiency in processing, organizing, and reflecting on information (DeYoung et al., 2003). Alongside openness to values and aesthetics, it has been found to be heavily associated with fluid intelligence, thereby showing cognitive flexibility, and not merely knowledge of facts (DeYoung et al., 2003; Wainwright et al., 2008). Nonetheless, researchers acknowledge that openness to ideas, or openness in general, cannot be used to measure intelligence because they describe a dimension of personality as opposed to intellectual ability (McCrae & John, 1992). High openness has also been frequently linked to education level, which has made the issue of using openness to determine intelligence or intellect controversial (Tesch & Cameron, 2003).

Since a characteristic of openness to ideas is the desire to seek and engage with different types of ideas, researchers have explored the relationship between creative thinking and openness. Silvia and colleagues (in press) had college students complete measures of the Big Five and measurements of creativity, including creative cognition and creative achievement. While the Big Five is intended to be a separate measure of personality, measurements of creativity also indirectly assess aspects of personality because the two areas of creativity, plasticity and stability, are a combination of the Big Five traits. Plasticity is composed of general openness to experience and extraversion to reflect the tendency "to explore and engage flexibly with novelty, in both behavior and cognition" (Silvia, Nusbaum, Berg, Martin, & O'Conner, in press). This area of creativity in particular is argued to be a higher-order factor

relevant to openness (Silvia et al., in press). Stability is a combination of the remaining Big Five traits, agreeableness, conscientiousness, and neuroticism/emotional stability, and reflects the tendency "to maintain stability and avoid disruption in emotional, social, and motivational domains" (Silvia et al., in press). They found that the students' levels of openness predicted their levels of creativity in several domains, ranging from arts and sciences to humanities (Silvia et al., in press). While there was a positive association though, there were a few concerns, all countering the assessment of creativity.

The first concern was regarding the plasticity dimension of creativity. Because it is composed of extraversion in addition to openness, it may not be entirely accurate. Some introverts have been found to be more open and creative than extraverts. It may be however, that introverts are more likely to have higher levels of stability due to the frequent associations between introversion and agreeableness, conscientiousness, and neuroticism. Whereas, they may have lower levels of plasticity due to the fact that it relies highly on extraversion. A second concern was regarding the finding on predicting creativity in several domains. High creativity and openness was found in students who had high interest in sciences, but if they had high interests in the arts. Other students however, such as those who exhibited interest in the math-science domain were not found to exhibit high levels of creativity. This may indicate a relation to openness to aesthetics, where intellect alone is not sufficient, and artistic creativity and interest is necessary. DeYoung, Quilty, and Peterson (2007) have proposed that openness may in fact be divided into an intellectual component and a more artistic component involving imagination, creativity, and aesthetics. Thus, while creativity can be a measure of plasticity and intellect, it may be a better measurement to use when assessing openness to aesthetics, instead of openness to ideas.

Biological correlates of openness to ideas. Little work has been done on the biological correlates of openness to ideas, with most research targeting biological correlates of intelligence to represent the facet. In one study by Duncan and colleagues (2000) PET scans of subjects were taken during tasks requiring the use of general intelligence (g). The tasks that required high use of g were found to activate the dorsolateral prefrontal cortex (PFC) as well as the dorsal

anterior cingulate cortex (ACC) (Duncan et al., 2000). It is proposed that this may be due to the PFCs role in cognitive permeability and the ACCs role in working with the PFC to monitor possible errors and reset task goals. It is also hypothesized that dopamine may regulate openness' cognitive permeability, as demonstrated in DeYoung et al.'s study on openness to actions where increased dopaminergic activation in the PFC was related to increased cognitive ability and flexibility (DeYoung et al., 2003).

Health and openness to ideas. Openness to ideas has generally been shown to provide numerous health benefits. A desire to engage in various activities and interests not only increases knowledge and keeps the brain significantly more active, but is also linked to many introspective and expressive behaviors that have the potential to raise self-esteem, flexibility, and life satisfaction (Tesch & Cameron, 2003; Wainwright et al., 2008). Individuals high on openness to ideas have been found to exhibit greater flexibility in dealing with life changes as well as changes in one's identity (Whitbourne, 1986).

Flexibility is linked to an analysis of identity exploration and a reflection on current events. Someone who is flexible exhibits the willingness to not only think about the changes that they're undergoing, but make reasonable alterations to their life (Whitbourne, 1986; Tesch & Cameron, 2003). When examining how openness to ideas was linked to personality and life change, Whitbourne found that through flexibility, individuals came to feel positive about themselves and that they had a sense of control over how to deal with different life events (Whitbourne, 1986). A previous study by Costa and McCrae (1980) had shown that there was a relationship between the occurrence of major life events and openness. Whitbourne's study came across a similar finding in which individuals who were more open were more likely to seek out experiences that increased knowledge, such as pursuing a higher education (Whitbourne, 1986).

Stephan (2009) examined the relationship between openness and life satisfaction in older adults and found that openness to ideas was positively correlated with self-reported life satisfaction, regardless of self-reported health and financial satisfaction. He concluded that greater life satisfaction was correlated with openness to ideas because open people typically search for opportunities for personal growth and reflect on their experiences. Opportunities pursued may include a wide range of intellectual and cultural activities that can enable individuals to enhance and maintain their cognitive abilities, which can promote higher life satisfaction. People high on openness to ideas not only exhibit higher life satisfaction, but also less stress. A study by Oswald and colleagues found that after taking a psychological stress test, individuals high on openness to ideas had lower cortisol responses than individuals low on the facet (Oswald et al., 2006). The study previously mentioned by Schneider and colleagues (2011) in the discussion on openness to actions, also felt that high openness to ideas predicted lower cortisol responses.

Openness to Values

Openness to values is defined as the extent to which individuals are resistant to or receptive of change. Similar to openness to ideas, openness to values has frequently been linked to determining flexibility of thought. Individuals who are high on openness to values question conventional norms and traditions and are more likely to reject unconventional principles. Thus, it is often studied in political psychology to analyze the personality differences between "conformists" and "non-conformists". Whereas conservative individuals are less adventurous, socially conforming, and prefer rigidity and structure in society, liberal individuals are more likely to support protests and revolutions in light of accepting change (McCrae, 1996). Some research has even found that low openness to values is linked to authoritarianism and a tendency towards aggression (McCrae, 1996). Research has also shown that cultural conservatism values more traditional work ethics and conventional female roles (Van Hiel & Mervielde, 2004). While openness to values encourages independent thinking, action and receptiveness of change, low openness to values encourages protecting stability and security (McCrae, 1996). Thus, it's believed that openness to values can predict whether one prefers change in aspects of her life or not.

Despite these positive findings, however some researchers have argued that there is a distinction between personality traits and moral values, and so it is conceptually invalid to claim that such differences reflect personality per se. As McCrae has argued, "traits describe what

people are like, but values refer to what people consider important" (McCrae, 1996). Traits are representative of enduring dispositions, whereas values serve as cognitive representations of enduring goals and guiding principles of how one prefers to live their life (McCrae, 1996). One study by Van Hiel and colleagues (2000) had shown that there was a relationship between conservative ideology and openness facets in Western Europeans. However, the one facet that exhibited no association was the values facet. Thus, it is not clear if psychological factors, no matter their ideological content, are linked to a liberal or conservative ideology (Hiel & Mervielde, 2004). Hiel & Mervielde argue that conservatism may be conceptualized differently across cultures, where in some cases they may be related to certain personality traits such as broadmindedness, but in other cultures, such as the Western European sample taken in Van Hiel et al.'s (2004) experiment may not.

Biological correlates of openness to values. Few investigations of biological correlations of openness to values have been performed. Those that have been done, however, have focused on cognitive flexibility, with the idea that greater cognitive flexibility is not only reflective of intellect, but also resistance to or acceptance of change. In one study by Kalbitzer and colleagues (2009), cerebral binding of plasma membrane serotonin transporter (5-HTT) was tested. Kalbitzer and colleagues (2009) hypothesized that potentiation of sertonergic transmission could effect cognitive flexibility. Using positron emission tomography (PET), it was found that subjects who had greater cerebral binding of the 5-HTT selective PET radioligand were lower on openness than subjects who had low levels of binding (Kalbitzer, Frokjaer, Erritzoe, Svarer Cumming, Nielsen, Hasemi, Baare, Madsen Hasselbalch, Kringelbach, Mortensen, & Knudsen, 2009). Interestingly, however, most of those who were low on binding were S-allele carriers. Many studies have shown increased vulnerability to various types of illness and disorders with the S-allele of the 5-HTT. However, in this case, having the S-allele seemed to have a beneficial effect. The researchers interpreted the association between low binding and high openness to change as reflecting greater cognitive flexibility. They believed that the lower binding promoted slower serotonin re-uptake at the plasma membrane so that there were higher extracellular serotonin levels (Kalbitzer et al.,

2009). Increased extracellular serotonin levels would lead to increased neural plasticity and responsiveness (Kalbitzer et al., 2009). Other studies have not followed up on these findings or proposed other biological correlates for being associated with openness to values.

Health and openness to values. Similar to investigations on biological correlates of the values facet, there has been little investigation on its relationship to health. Most studies however have demonstrated that being high on openness to values can be harmful to one's health. Jonassaint and colleagues (2007) found that individuals who were high on values were at increased risk for cardiac deaths and elevated mortality, compared to individuals who were low on values. The researchers did not interpret this result. However, it is possible that the desire to stick with familiarity and avoid accepting change may lead to a Type A personality that is frequently discussed in medical literature. Another study on the relationship between health and openness to values found that current smokers had significantly higher openness to values scores than non-smokers in the study (Terracciano & Costa, Jr., 2004). This is contrary to the belief that increased smoking would be linked to openness to change. Since findings have produced mixed results, further investigation in the area of openness to values and health needs to be done.

Openness to Aesthetics

While openness to actions, ideas, and values focus on the interaction between cognition and action, the remaining three facets, aesthetics, fantasy, and feelings focus on emotional factors. Openness to aesthetics, the first of the three more emotional facets, is described as the appreciation of different art forms. While individuals high on openness to aesthetics tend to enjoy and appreciate the arts, the facet does not apply to appreciation for a particular kind of art form (Tellegen & Atkinson, 1974). Aesthetic sensitivity carries with it a particular form of intellect that combines intelligence, creativity, imagination, and perception (McCrae & John, 1992). DeYoung et al.'s experiment, mentioned earlier, on dopamine projections and cognitive performance had examined the role of aesthetics. While the main finding of the experiment was that openness to actions was the facet least linked to cognitive flexibility and

intelligence, the role of openness to aesthetics proved to be intimately linked to cognitive flexibility and intelligence (DeYoung et al., 2003). In terms of creativity and imagination, some studies have found that higher scores on aesthetics have been significantly linked to creative activities (Griffin & McDermott, 1998).

What makes the aesthetics facet distinct from the other facets though is not its role in intellect or even creativity, but its appreciation of experience for its own sake (Bergeman, Chipuer, Plomin, Pedersen, McClearn, Nesselroade, Costa, Jr., & McCrae, 1993). While most assessments of openness to aesthetics have relied on questionnaires, some studies have looked at aesthetic markers as being a universal marker of openness to experience, which have been particularly unique to the openness trait (McCrae, 2007). Aesthetic chills are emotional responses to experiences of beauty (McCrae, 2007). The emotion experienced during aesthetic chills is similar to those felt during the emotion of elevation, in which there is a feeling of awe and deep appreciation. Thus, while the chills experienced can be similar to chills of surprise, they are not unpleasant (McCrae, 2007). They also differ from the chills of excitement that come with risk-taking, which are likely to be more apparent with high openness to actions (McCrae, 2007).

While everyone has the capacity to experience aesthetic chills, individuals who are more prone to experiencing them are those who exhibit high emotional responsivity and sensitivity. They are the type of people who pay greater attention to all kinds of stimuli, internal and external, as argued to be linked to a temperament known as Orienting sensitivity (McCrae, 2007; Evans & Rothbart, 2008). They are also the type to enjoy connecting with experiences by absorbing them and engaging with them emotionally (Tellegen & Atkinson, 1974). Thus, people high on the aesthetics facet frequently experience opposing emotions simultaneously and at high intensities, though not at the level of a clinical mood disorder (McCrae, 2007). As a result of the strong emotional component, they also tend to be high on the facets of fantasy and especially, feelings. In addition, these facets have been found to be significantly high in individuals who are low on extraversion, but high on openness to experience (DeYoung et al., 2003). Thus, the proactive seeking of breadth and depth of experiences need not merely be expressed through high scores on the actions facet, but

also on the more emotional facets (DeYoung et al., 2003).

Health and openness to aesthetics. Minimal research has been done on the health benefits of being high on openness to aesthetics. However, the few studies that have been done have examined the link between the aesthetics facet and artistic creativity because artistic creativity has been found to correlate with mood disorders (Wolfestein & Trull, 1997). In one study that explored openness to aesthetics, researchers found that the facet was most strongly related to depression (Wolfestein & Trull, 1997). As most studies have previously demonstrated, there is a significant and common association between and neuroticism. In this study, the researchers found that individuals higher on openness to aesthetics were less extraverted, more neurotic, and scored higher on depression scores (Wolfstein & Trull, 1997). The implications of these findings are not completely clear because while individuals may engage in the arts out of depression, engaging in arts may also decrease their feelings of depression. The findings on aesthetics, introversion, and neuroticism are also unclear. It is well known though that introverts are often more emotional, regardless of whether they express it or not.

Openness to Fantasy

Like openness to aesthetics, openness to fantasy is also strongly related to a person's emotional disposition. As the name of this facet implies, it encompasses the tendency to engage in fantasizing, which not only involves the development of a mental picture, but also engages high levels of creativity and several emotions. While most research on fantasizing has found that it can lead to increased risk for depression, the research also suggests that fantasizing can have both positive and negative effects, depending on the emotional investment that one puts in them, the types of fantasies that one has, and whether or not the fantasies effect everyday behavior.

Health and openness to fantasy. Returning to Wolfestein & Trull's study, as discussed in the section on openness to aesthetics, openness to fantasy is also linked to depression (Wolfstein & Trull, 1997). Wolfestein & Trull (1997) predicted that openness to fantasy could have a positive

relation with mental health since openness and self-actualization are closely linked. However, they also kept in mind that a private self-focus, a construct highly linked to openness to fantasy and openness to feelings, could be linked to depression. The Self-Awareness Theory of Depression argues that when individuals experience a type of loss or failure then they face a challenging time in dealing with the differences between their actual and ideal states (Pyszczynksi, Holt, & Greenberg, 1987). This is where fantasizing can potentially play a significant role and lead to increased risk for depression and negative emotions (Wolfstein & Trull, 1997). Keeping this theory in mind, the researchers anticipated the possibility of individuals high on openness to fantasy experiencing a difficult time in dealing with their actual and ideal states (Wolfstein & Trull, 1997). While the results demonstrated that openness to fantasy was linked to depression, there was a greater link to depression if the subject was extraverted as opposed to introverted.

Another study by Garcia et al. (2005) produced the same findings, where fantasizing was more detrimental for the mental health of an extravert than an introvert. Wolfstein and Trull (1997) have hypothesized that extraverts may be more susceptible to having negative fantasies that deal with failing social relationships, whereas introverts may not engage as much in socially based fantasies, and this may explain the result. A study by Carillo and colleagues (2001) has also examined the role of openness to fantasy in predicting depression and came across similar findings as Wolfstein and Trull. In addition to Wolfstein and Trull's findings however, they also examined gender differences and found that women scored higher in fantasy than men, and were thus more susceptible to depression than men (Carillo et al., 2001).

In a previous study by Carrillo and colleagues it was found that fantasy predicted more than just depression, but also "neuroticism, family maladjustment, personality disorders, and a lack of positive affect" (Carrillo, Rojo, Sanchez-Bernados, & Avia, 1998). It may be that individuals who fantasize are more likely to do so because they are unhappy with reality and fantasizing gives them a sense of control over their own inner reality. It has been proposed that overly positive fantasies about the future can actually have negative effects on one's health by potentially suppressing motivation and action, thus increasing one's vulnerability to depression

(Carrillo, Rojo, & Staats,1996). However, it is also a possibility that positive fantasies can provide one with ambition to act and change their circumstances, so long as they are within realistic bounds and are not impossible. In addressing the possibility of openness facets, including fantasy, leading to increased risk for depressing, McCrae and Costa argue that openness is not meant to lead to either positive mental health or maladjustment (Carrillo et al., 2001). They propose that an individual high on openness is able to experience and reflect on several positive and negative experiences. By undergoing a wide variety of experiences, an individual high on openness is able to find a balance of positive and negative reactions to their experiences (Carrillo et al., 2001).

In the area of physical health, there have been studies on imagination and disease contraction. One study by Sherman and colleagues (1985) examined how perceived likelihood of contracting a disease can be influenced by imagining contraction of the disease. Some participants were asked to imagine easy-to-imagine symptoms and others were asked to imagine hard-to-imagine symptoms. Afterwards, participants were asked to rate how easy it was to imagine the symptoms in the assigned condition and how likely they were to actually contract the disease (Sherman, Cialdini, Schwartzman, & Reynolds, 1985). Sherman and colleagues (1985) found that judgment of imagination as being easy or difficult reflected participants' estimates of likelihood of contracting the disease. The subjects who rated the disease as being easy-to-imagine had judged the disease as more likely to occur than subjects who rated the disease as being hard-to-imagine (Sherman et al., 1985). Perhaps for some, fantasizing is a source of motivation for changing the way they perceive a future negative event, while for others, it may be a source of depression because they might be more inclined to have negative fantasies. Regardless of whether the effects of fantasizing are positive or negative though, the influence of fantasizing appears to have significant implications because it demonstrates the power of appraisal, where thinking about a situation can influence one's perception of an event.

Openness to Feelings

Of all six facets, openness to feelings is recognized by many researchers as being the

most complex (Coan, 1972; McCrae, 2007). Individuals high on openness to feelings highly value emotions and are thus, more sensitive to emotional experiences. They also experience most emotions more intensely than individuals who are lower on the facet. Greater levels of sensitivity and intensity of emotional experience has often been noted as being potentially harmful to one's health. In certain cases, experiencing a wide range of emotions can result in frustration because an individual may be confused and overwhelmed with their emotions (McCrae, 2007). The inability to make sense of the mixed emotional states that one is experiencing may be even worse for people who frequently suppress their emotions. A lack of coherence of emotional states has the potential to lead to difficulty in effective emotion regulation. Unsurprisingly, research shows that high openness to feelings has been linked to higher experience of anxiety and neuroticism (Garcia et al., 2005). Research also shows that women tend to be higher on openness to feelings than men and that this is linked to the higher rates of anxiety and mood disorders in women (Coan, 1972). Another large body of research though proposes that emotional awareness can actually be beneficial and reduce health problems, including depression.

Biological correlates of openness to feelings. Research on the biological correlates of openness to feelings has generally been linked to understanding the role they play in disease and psychiatric disorders. In one study it was found that a functional polymorphism of the glutathione peroxidase 1 (GPX1) gene was linked to openness to feelings, and even openness to experience in general cell (Matsuzawa, Hashimoto, Shimizu, Fujisaki, & Iyo, 2005). Glutathione peroxidases protect cells from oxidative damage by reducing hydrogen peroxide present in and around the cell (Matsuzawa et al., 2005). The three possible genotypes are Pro198Leu, Leu198Leu, and Pro198Pro. Unlike the Pro allele, previous studies have shown that the Leu allele is less responsive to added selenium and may lead to functional consequences. In this study, individuals who had the Pro198Pro allele were higher on openness to experience, and the feelings facet in particular, than any genotypes with the Leu allele (Matsuzawa et al., 2005). Matsuzawa and colleagues argue that this may have significant implications for risk of neuropsychiatric diseases and mood disorders,

especially since the Pro allele demonstrates a protective effect and is more beneficial than the Leu allele.

Other studies have looked at the brain-derived neurotrophic factor (BDNF) gene Val66Met polymorphism, particularly since it is associated with the pathophysiology of mood disorder, as is openness to feelings (Matsuzawa et al., 2005). BDNF has been heavily investigated for its role in depression, and antidepressants have aimed to increase BDNF transcription to cure symptoms of the disorder (Sen, Nesse, Stolenberg, Li, Gleibernman, Chakarvarti, Weber, & Baurmeister, 2003). Other reasons for investigating the relationship between BDNF and depression come from research on depression, hippocampal size and neurogenesis. Individuals who are depressed have smaller hippocampi and experience decreased neurogenesis. BDNF is a crucial factor in neurogenesis because it affects neuronal differentiation during development and plays a crucial role during adulthood in synaptic plasticity and neuronal survival (Sen et al., 2003). In a study by Sen and colleagues it was found that the Val allele was linked to an increase for depression and that the Met allele had a protective effect. This is because the Mel allele of BDNF produces higher activity and better processing of BDNF. Individuals who were high on openness to feelings were more likely to have a homozygous or heterozygous genotype including the Mel allele (Sen et al., 2003).

One study by Jonassaint and colleagues examined how the various facets of openness to experience were linked to inflammation. They found that the feelings facet had the most significant effect. Similar to the above findings they also found that higher openness to feelings had beneficial effects. Individuals higher on openness were found to have lower mean C-reactive protein (CRP) levels. They proposed that this was related to high emotional awareness and can have a protective effect against disease. This is unlike low emotional awareness, which they suggested can lead to excessive stress-induced sympathetic activation and affect circulating inflammatory markers (Jonassaint, Boyle, Kuhn, Siegler, Copeland, & Williams, 2010). The challenge of identifying feelings and failing to allow oneself to fully experience them has been linked to greater physiological arousal (Jonassaint et al., 2010). Thus, low emotional awareness may lead to increased risk of inflammation and

cardiovascular disease (CVD). The researchers also proposed that low emotional awareness can lead to negative affective states that have the potential to elevate CRP.

Another study by Jonassaint and his colleagues (Jonassaint, Boyle, Williams, Mark, Siegler, & Barefoot, 2007) also came to the same conclusion: that low openness to feelings can increase risk for illness. They found that low openness to feelings led to cardiac trends and that there was a protective effect of scoring higher on this facet (Jonassaint, 2007). They propose that this is due to low emotional awareness and alexithymia-blunted affect. Low emotional awareness can lead to increased physiological arousal and predispose individuals to behaviors like social isolation and substance abuse and increase the experience of affective states including "depression, somatic complaints, and anxiety" (Jonassaint, 2007). Alexithymia-blunted affect can also lead to diseases through "physiological, behavioral, social, or cognitive mechanisms". What is interesting to note though is that the individuals high on openness had experienced positive and negative emotions more intensely, which can also be another risk in itself.

Health and openness to feelings. Some studies have examined how various types of emotion regulation and appraisal can influence health. Such findings are significant because both extreme emotional suppression and extreme emotional expression have been linked to emotion in disease onset and progression, affecting the endocrine, immune, and autonomic systems (Gross, 1989). For instance, in one study it was found that frequent suppression of negative emotions predicted increased risk for myocardial infarction (MI) (Denollet & Brutsaertm, 1998). In another study it was found that individuals who suppressed their emotional thoughts were also more susceptible to general illness (Petrie, Booth, & Pennebaker, 1998). They showed a significant decrease in CD3T lymphocyte levels, CD8 (T suppressor) cells, and total lymphocyte numbers (Petrie et al., 1998). They also had poorer natural (NK) cell activity. However, the researchers also found that those who had done emotional writing and were high on openness to aesthetics had increased levels of circulating CD4 (T helper cells) and the total number of lymphocytes (Petrie et al., 1998). This protective effect has been seen in other studies, where writing about thoughts and emotions has proven to be

therapeutic for individuals (Cameron & Nicholls, 1998). Writing always leads individuals to take different approaches to their experiences and think of ways to resolve potential conflicts. The worry though is that while some individuals might use writing to continuously ruminate on stressful events instead of taking the time to develop effective coping strategies (Cameron & Nicholls, 1998). Thus, it is not necessarily the intensity of emotional experience that can be costly, but the way one deals with their emotions.

Whether or not openness to feelings is beneficial seems to depend significantly on appraisal since appraisals guide emotional experience and action (Schneider, 2008). As an indirect measurement of the effects of appraisal, researchers have examined how and if stress responses are linked to openness to feelings. This is because the way one deals with stress involves changes in appraisals, emotion, and task performance (Schneider, 2008). Appraisal describes how one deals and interacts with a situation in the environment (Lazarus, 1999). Primary appraisal begins with an evaluation of whether or not a stressor is relevant to an individual. A secondary appraisal evaluates the resources that are available for coping with the stressor. If an individual thinks of the situation as a challenge that they are capable of overcoming, it might reasonably be predicted that they would experience a greater level of confidence in their ability to cope with the situation. However, if an individual sees the situation as a threat then they are likely to feel overwhelmed and fail to deal with it effectively (Lazarus, 1999).

In a study on the effects of appraisal on affect and performance, Schneider (2008) found that challenge appraisals led to greater positive affect and that threat appraisals led to greater negative affect. She also found that openness in particular was what drove this effect and that extraversion, which is often expected to be related to appraisals, was not (Schneider, 2008). When presenting subjects with a task, they found that performance was also associated with openness, with those higher on openness to feelings performing better (Schneider, 2008). A study specifically looking at stress regulation and openness found that higher openness to feelings was also linked to increased positive affect (Williams, Rau, Cribbet, & Gunn, 2009).

In terms of chronic illness onset and progression, higher openness to feelings has been linked to better health outcomes. It has been linked to decreased cardiac death and

lower all-cause mortality (Williams et al., 2009). Researchers have also found that individuals higher on openness have blunted cortisol responses and experience higher parasympathetic activity and decreased sympathetic activity (Williams et al., 2009). The belief is that this results from effective reappraisal of negative emotional stimuli and the possibility that more open people are higher on mindfulness when it comes to emotion labeling. Thus, even when confronted with difficult situations, individuals who are high on openness to feelings may engage in more adaptive and flexible coping mechanisms (Williams et al., 2009). This type of mindfulness may even extend to coherence of narratives from psychotherapy experiences, allowing individuals to engage in meaningful reflection on their experiences. Some evidence for this is linked to prefrontal cortex functioning, which is linked to general cognitive abilities, as discussed in the section on openness to actions. In addition to playing a role in cognitive abilities, the PFC also plays a role in stress-dampening self-regulatory activity. Thus, the PFC's role in executive functions, which is typically more active in people high on openness, also includes inhibitory control.

The ability for high-open individuals to engage in more effective emotion regulation and see things as being more rewarding than threatening has extended to research on openness and life satisfaction (Williams et al., 2009). Stephan has examined the relationship between openness and life satisfaction in older adults. He found that openness to feelings was positively correlated with self-reported life satisfaction, while controlling for health and financial satisfaction (Stephan, 2009). When it came to the benefits of openness to feelings, Stephan believed that this resulted from open individuals adjusting how they felt in response to experiencing new ideas or situations (Stephan, 2009). Thus, the inability to adjust or cope effectively can in fact lead to greater vulnerability to the effects of stress for low-open individuals and provide a protective effect for those who are high on openness to feelings (Williams et al., 2009).

Openness to Experience: A Good Thing?

Personality plays a key role in our lives because of the significant impact it has on the decisions we make such as the activities that we choose to pursue and the way we choose to perceive and deal with events in our lives. The pursuit of understanding what the core elements of personality are remains as different theories continue to arise. As the field of personality psychology continues to grow though, some aspects of personality gain wide acceptance. The Five Factor Model is an instance of this. Its comprehensive integration of past research and personality scales has helped further the understanding of personality. Even though there are multiple questionnaires that adopt the theory, they adopt the same five main traits: extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience, and have been to produce similar results. Only recently has application of this model been extended to fields beyond personality psychology, in hopes of gaining a more extensive understanding of human mind and behavior. Instances of this can be seen in fields exploring health such as psychology and medicine, where fields that once appeared to be unrelated are now relying significantly on each other.

To better understand physical and mental health, researchers have begun to acknowledge the important role of personality and thus, have adopted the popular Five Factor Model into many of their studies. Researchers have stumbled across several correlations between the Big Five traits and aspects of health such as health perception, health behavior, risk for disease, and disease progression. While the traits of extraversion, neuroticism, agreeableness, and conscientiousness have produced consistent results across numerous studies, findings on the influence of openness to experience on health have remained mixed and heavily disputed. In this review, we proposed that the mixed findings may have been due to the complexity of the trait and thus, proposed that the trait be examined in more depth, instead of being examined broadly. Openness to experience, the most controversial of the Big Five, can describe a wide array of people such as those who appreciate the arts, those who enjoy traveling and trying new foods, and those who love to fantasize. Since people can be open in a myriad of ways, we suggested that a detailed look at individual levels of the main Big Five traits might give a clearer understanding of openness to experience as well as its implications for health.

To gain an accurate understanding of the actions facet, we first examined the similarities

and differences between openness to actions, extraversion, and sensation-seeking, finding that the actions facet not only had a strong behavioral component, like the other two personality constructs, but also a strong cognitive component. We identified some potential biological correlates including norepinephrine, MAO, and variations in dopamine receptors (Cloninger, 2000). In our investigation of its effect on health, we found that scoring high on the facet had the potential to be beneficial to one's health because it was associated with a greater experience of positive emotion (Carrilo et al., 2001). This was also true for those who were diagnosed with an illness. Openness to actions appeared to form a buffer against depression because it encouraged one to pursue activities that brought them joy, regardless of their current state (Salovey et al., 2000). At the same time, openness to actions has the potential to be costly to one's health. One who pursues dangerous activities such as substance abuse may be experiencing a lot of positive emotion, but also face many health risks (Booth-Kewley & Vickers, Jr., 1994). Future studies should identify what other unhealthy behaviors may be linked to openness to actions besides substance abuse.

In our analysis of the openness to ideas facet, we found that individuals scoring high on the facet felt more positive emotion and were more satisfied with their life (Stephan, 2009). The suggested reason for this is that the brain is kept active and that individuals are more flexible in thinking, as supported by some studies on the role of dopamine in cognitive flexibility (Duncan et al., 2000). Thus, not only do individuals get joy out of increasing their knowledge, but they are also able to better cope with challenges they encounter because they make the effort to think about them positively (Whitbourne, 1986; Tesch & Cameron, 2003). While these findings are linked to cognitive flexibility, researchers must be careful not to confuse the facet with intellect or intelligence. Much of the debate that has come about regarding what openness is as a trait has come about from the cognitive characteristics related to this facet.

Unlike our findings on the actions and ideas facets, we did not come across positive outcomes linked to the values facets. The few studies that have been done on its role in health have found that it is linked to increased risk for cardiac deaths as well as elevated mortality (Jonassiant et al., 2007). No interpretation of these results was provided by the researchers

behind this study. Other studies however, indicated that this may be linked to decreased cognitive flexibility, which is effected by sertonergic transmission. A primary concern regarding the openness to values facet is where to draw the line between one's values and one's personality. It is not entirely clear how holding onto one's traditional values can have a negative influence on their health, especially if they are content with those values and use them to structure aspects of their life.

Similar to the values facet, we found minimal research on the role of the aesthetics facet in health. However, we reviewed related research similar to the aesthetics facet. Research on artistic creativity had shown that it was sometimes associated with mood disorders. In one study, it was found that just like increased artistic creativity, a higher score on the aesthetics facet was linked to increased risk for depression (Wolfestein & Trull, 1997). While we did not find any studies that looked at biological correlates of the aesthetics facet, perhaps it would be useful to look at the role of BDNF and the GPX1 gene, which have been found to increase risk for depression. Biological correlates aside though, it is challenging to determine whether engagement of the arts is a result of depression or if it helps to decrease depression. At the same time, the aesthetics facet is not entirely equivalent to artistic creativity since the aesthetics facet need not entail engaging in the arts, but simply appreciating the arts.

Just as we did not come across any findings that examined biological correlates of openness to aesthetics, we also did not find any that were related to the facet of fantasy. We did however, come across similar findings to the aesthetics facet in that the fantasy fact was also linked to depression (Carrillo et al., 1996; Wolfestein & Trull, 1997). Some studies indicated between-group differences though, finding that extraverts high on the fantasy facet were more susceptible to negative outcomes than introverts. The researchers behind the study proposed that this may have been due to differences in fantasy content, namely that extraverts may have more negative socially-based fantasies. Further investigation should be done on how extraversion might play a role in fantasizing and more generally, how the content of one's fantasies can influence their health and well-being. Just as we proposed in our discussion on the facet, it may be that for some fantasizing is a

source of motivation and hope, while for others it may be a source of depression and loss of hope.

The last facet we examined was the feelings facet, which was the most unclear in its implications for health relative to the other five facets. On one hand, we found that some of its biological correlates such as the GPX1 gene and the BDNF gene increased risk for neuropsychiatric diseases and mood disorders. On the other hand, we found that low scores on the facet were linked to decreased risk of inflammation and contraction of cardiovascular disease (Jonassaint et al., 2010) as well as decreased cardiac death and greater life satisfaction (Williams et al., 2009). It may be beneficial to be aware of one's emotions, but emotional awareness does not automatically imply that one knows how or even wants to deal with her emotions. For instance, one may ruminate continuously or suppress their emotions, both of which can be harmful to one's health (Gross, 1989). In this case, it appears to be more beneficial to be less emotionally aware. Future studies need to investigate the relationship between emotional awareness and emotion regulation, and ultimately, how this relationship impacts one's health.

When examining the effects of openness to experience on health, our analysis of each facet enabled us to understand why the trait of openness was so complex. We suggest that a facet approach be taken by empirical researchers and that they use the Big Five questionnaires that contain facets as opposed to using the more condensed questionnaires because the questionnaires with facets are more detailed. Beyond the obstacles that come with understanding the role of each facet level in health though, there is another obstacle that arises with the independence of the trait itself. Openness to experience has been continuously challenged for not being an independent trait even though psychometric tests have demonstrated that it is a very distinct trait (Garcia et al., 2005). For instance, he actions facet is frequently linked to extraversion, the ideas facet is often correlated with agreeableness, and the values facet tends to be related to neuroticism.

Related to the independence of openness is the independence of the other Big Five traits themselves. The Five Factor Model may be a comprehensive and widely used model to assess personality, but it is not the only way to assess personality. It has been criticized for being too broad and failing to take into account aspects of personality that go beyond the main five factors (McCrae & John, 1992). It has also been criticized for its questionnaire approach, which puts a limitation on personality. There is the risk of bias on the part of the participant because individuals are more likely to report positive traits on questionnaires as opposed to negative traits. At the same time, personality is dynamic and operates through a variety of situations. Even though certain patterns are consistent, questionnaires cannot capture this dynamic and can only provide one side of the picture. Detailed approaches such as narratives should be examined in addition to questionnaires because they are likely to offer a different perspective in understanding personality. Through a narrative, one is able to provide descriptions that give explanations for why they may have acted or felt a certain way. Unfortunately, a structured questionnaire with questions that have been written by researchers does not offer free responses.

If researchers hope to understand the role of openness to experience in health, it is imperative that they look at how the individual aspects of the trait influence health and attempt to separate openness from the other Big Five traits when doing so. More importantly however, they should consider multiple ways of assessing personality because the Five Factor Model, as any personality model, will have its limitations. Given that the field of personality psychology is complex and ever-changing though, researchers should keep in mind that the results will be ever-changing as well.

References

Aluja, A. et al. (2002). Relationships among extraversion, openness to experience, and sensation seeking. *Personality and Individual Differences 35:671-680.*

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social and General Psychology Monographs 132:455-929.*

Bergeman, C.S. et al. (1993). Genetic and Environmental Effects on Openness to Experience, Agreeableness, and Conscientiousness: An Adoption/Twin Study. *Journal of Personality 61:2.*

Bermudez, J. (1999). Personality and Health-Protective Behaviour. *European Journal of Personality, 13,* 83-103.

Booth-Kewley, S.B., & Vickers, Jr., R.R. (1994). Associations between Major Domains of Personality and Health Behavior. *Journal of Personality 62:3.*

Boyce, C.J., & Wood, A.M. (2011). Personality Prior to Disability Determines Adaptation: Agreeable Individuals Recover Lost Life Satisfaction Faster and More Completely. *Psychological Science, 22(11),* 1397-1402.

Cameron, L.D., & Nicholls, G. (1998). Expression of Stressful Experiences Through Writing: Effects of a Self-Regulation Manipulation for Pessimists and Optimists. *Health Psychology 171):84-92.*

Capitanio et al. (2008). Personality and serotonin transporter genotype interact with social context to affect immunity and viral set-point in simian immunodeficiency virus disease. *Brain, Behavior, and Immunity, 22 (5),* 676-689.

Carrillo, J.M. et al. (2001). Openness to Experience and Depression. *European Journal of Psychological Assessment 17(2):130-136.*

Carrillo, J.M. et al. (1998). *Around openness to Experience: Its role in psychological adjustment.* Communication presented at the 9th European Conference on Personality. Surrey, England, July.

Carrillo, J.M. et al. (1996). Vulnerable personality in depression: Investigating communality in the search for unification. *European Journal of Psychological Assessment* 12:202-211.

Cloninger, C.R. (2000). Biology of personality dimensions *Current Opinion in Psychiatry 13:611-616.*

Coan, R.W. (1972). Measurable Components of Openness to Experience. *Journal of Consulting and Clinical Psychology 39(2):346.*

Costa, P.T., & McCrae, R.R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology 38:668-678.*

Denollet J. & Brutsaertm, D.L. (1998). Personality, Disease Severity, and the Risk of Long-term cardiac Events in Patients With a Decreased Ejection Fraction After Myocardial Infarction. *Circulation 97:167-173.*

Depue, R.A., & Collins, P.F. (1999). Neurobiology of the Structure of Personality: Dopamine, Facilitation of Incentive Motivation, and Extraversion. *Behavioral and Brain Sciences 22(3):491-517.*

DeYoung, C.G. et al. (2007). Between facets and domains: 10 Aspects of the Big Five. *Journal of Personality and Social Psychology 93:880-896.*

DeYoung, C.G., et al. (2003). Sources of Openness/Intellect: Cognitive and Neuropsychological Correlates of the Fifth Factor of Personality. *Journal of Personality 73:4.*

Dua, J. (1990). Assessment of positive and negative affect as a result of thoughts and real life experiences. *Behavior Change 7:62-65.*

Duncan, J. et al. (2000). A Neural Basis for General Intelligence. *Science Magazine 289:457-460.*

Epstein, R.P et al. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking. *Nature genetics 12:78.*

Evans, D.E., & Rothbard, M.K. (2008). Temperamental sensitivity: Two constructs or one? *Personality and Individual Differences 44(1):108-118.*

Friedman, H.S., & Booth-Kewley, S. (1987). The "Disease-Prone Personality": A Meta-Analytic View of the Construct. *American Psychologist 42(6):549-555.*

Garcia, L.F. et al. (2005). Is Openness to Experience an Independent Personality Dimension?: Convergent and Discriminant Validity of the Openness Domain and its NEO-PI-R Facets. *Journal of Individual Differences 26(3).*

Gerra, G. et al. (1999). Neurotransmitters, Neuroendocrine Correlates of Sensation-Seeking Temperament in Normal Humans. *Neuropsychobiology 39:207-213.*

Griffin, M., & McDermott, M.R. (1998). Exploring a tripartite relationship between rebelliousness, openness to experience and creativity. *Social Behavior and Personality 26(4):347-356.*

Gross, J. (1989). Emotional Expression in Cancer Onset and Progression. *Soc. Sci. Med. 28:1239-1248.*

Jerram, K.L., & Coleman, P.G. (1999). The big five personality traits and reporting of health problems and health behavior in old age. *British Journal of Health Psychology 4(2):181-192.*

Jonassaint C.R. et al. (2010). Personality and Inflammation: The Protective Effect of Openness to Experience. *Ethn Dis. 20(1):11-14.*

Jonassaint, CR. et al. (2007). Facets of Openness predict Mortality in Patients with Cardiac Disease. *Psychosomatic Medicine 69:319-322.*

Kalbitzer, J. et al. (2009). The personality trait openness is related to cerebral 5-HTT levels *NeuroImage 45:208-285.*

Kaasinen, V. et al. (2001). Personality traits and brain dopaminergic function in Parkinson's disease. *PNAS 98:23.*

Lazarus, R.S. (1999). *Stress and emotion: A new synthesis.* New York: Springer.

Matsuzawa, D. et al. (2005). Functional Polymoprhism of the Glutathione Peroxidase 1 Gene Is Associated with Personality Traits in Healthy Subjects. *Neuropsychobiology 52:68-70.*

McCrae, R.R. (2007). Aesthetic Chills as a Universal Marker of Openness to Experience. *Motiv Emot 31:5-11.*

McCrae, R.R. (1996). Social Consequences of Experiential Openness. *Psychological Bulletin 120(3):323-337.*

McCrae, R.R, & Costa, P. (1997). Personality trait structures as a human universal. *American Psychologist, 52:509-515.*

McCrae, R.R., & John, O.P. (1992). An Introduction to the Five-Factor Model and Its Applications *Journal of Personality 60(2):175-215*

Menza, M.A., & Mark, M.H., (1994). Parkinson's Disease and Depression: The Relationship to Disability and Personality. *The Journal of Neuropsychiatry and Clinical Neuroscience 6:165-169.*

Okuyama, Y. et al (2000). Identification of a polymorphism in the promoter region of DRD4 associated with the human novelty seeking personality trait. *Molecular Psychiatry 5:64-69.*

Oswald, L.M. et al. (2006). Relationship between Cortisol Responses to Stress and Personality. *Neuropsyhopharmacology 31:1583-1591.*

Panksepp, J. (1998). Affective neuroscience: The foundtions of human and animal emotion. New York: Oxford University Press.

Petrie, K.J. et al. (1998). The Immunological Effects of Thought Suppression. *Journal of Personality and Social Psychology 75(5):1264-1272.*

Pyszczynski, T. et al. (1987). Depression, self-focused attention and expectancies for future positive and negative events for self and others. *Journal of Personality and Social Psychology 52 994-1001.*

Reuter, M., & Henning, J. (2003). Association of the functional catechol-O-methyltransferase VALI58MET polymorphism with the personality trait of extraversion. *NeuroReport 16(10).*

Salovey, P. et al. (2000). Emotional States and Physical Health. *American Psychologist 55(1):110-121.*

Samochowiec, J. et al. (2004). Polymorphisms in the Serotonin Transporter and Monoamine Oxidase A Genes and Their Relationship to Personality Traits Measured by the Temperament and Character Inventory and NEO Five-Factor Inventory in Healthy Volunteers. *Neuropsychobiology 50:174-181.*

Schaller, M., & Murray, D. R. (2011). Infectious disease and the creation of culture. In M. Gelfand, C.-y. Chiu, & Y.-y. Hong (Eds.), *Advances in Culture and Psychology (Vol. 1,* pp. 99-151). New York: Oxford University Press.

Schneider, T.R. (2008). Evaluations of stressful transactions: what's in an appraisal? *Stress and Health 24:151-158.*

Schneider, T.R. et al. (2011). The Influence of Neuroticism, Extraversion and Openness on Stress Responses. *Stress and Health.*

Sen, S. et al. (2003). A BDNF Coding Variant Is Associated with the NEO Personality Inventory Domain Neuroticism, a Risk Factor for Depression. *Neuropsychopharmacology 28:397-401.*

Silvia, P.J. et al. (in press). Openness to experience, plasticity, and creativity: Exploring lower-order, high-order and interactive effects. *Journal of Research in Personality.*

Sherman, S.J. et al (1985). Imagining Can Heighten or Lower the Perceived Likelihood of Contracting a Disease: The Mediating Effect of Ease of Imagery. *Pers Soc Psychol Bull 11:118.*

Sloan et al. (2008). Social temperament and lymph node innervation. *Brain, Behavior, and Immunity, 22 (5),* 717-726.

Stephan, Y. (2009). Openness to experience and active older adults' life satisfaction: A trait and facet-level analysis. *Personality and Individual Differences 47:637:641.*

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology 83:268-277.*

Terracciano, A., & Costa, Jr., P.T. (2004). Smoking and the Five-Factor Model of Personality. *Addiction 99(4):472-481.*

Tesch, S.A., & Cameron, K.A. (2003). Openness to Experience and Development of Adult Identity. *Journal of Personality 55(4):615-630.*

Van Hiel, A., & Mervielde, I. (2004). Openness to Experience and Boundaries in the Mind: Relationships with Cultural and Economic Conservative Beliefs. *Journal of Personality 72:4.*

Van Hiel, A. et al. (2000). The relationship between Openness to Experience and political ideology. *Personality and Individual Differences 28(4):741-751.*

Wainwright, M.A. et al. (2008). Genetic Covariation Among Facets of Openness to Experience and General Cognitive Ability. *Twin Research and Human Genetics 11(3):275-286.*

Whitbourne, S.K. (1986). Openness to Experience, Identity Flexibility, and Life Change in Adults. *Journal of Personality and Social Psychology 50(1):163-168.*

Williams, P.G. et al. (2009). Openness to Experience and stress regulation. *Journal of Research in Personality 43:777-784.*

Wolfenstein, M. & Trullm, T.J. (1997). Depression and Openness to Experience. *Journal of Personality Assessment 69(3): 614-632.*

Zuckerman, M. (1995). Good and Bad Humors: Biochemical Bases of Personality and Its Disorders. *Psychological Science 6(6):325-332*

Zuckerman, M. (1979). *Sensation seeking: beyond the optimal level of arousal.* L. Erlbaum Associates.

Morning versus Evening Circadian Tendencies: Suicidality and Risk Propensity during the Transition to Adulthood

Gerard L. Markham
University of Rochester

Ellie McGlinchey, MA
University of California, Berkeley

Kerrie Hein, MA
University of California, Berkeley

Allison Harvey, PhD
University of California, Berkeley

People in the transitional period to adulthood who exhibit a circadian tendency toward eveningness (E-types) follow a delayed sleep schedule, increasing activity later in the day and both going to sleep and getting up later compared to morning-types (M-type). The circadian tendency toward eveningness during this transitional period has been associated with a wide range of adverse effects, including poor self-regulation, increased rates of substance use, impulsivity, depression, anxiety and emotional instability and more aggressive and antisocial behavior. We utilized public data from the National Longitudinal Study of Adolescent Health and analyzed a subsample of participants in the transitional period to adulthood aged 18-21 (n=2143). Participants were interviewed at four waves over the course of 15 years and were analyzed on indices of sleep chronotype, risk taking propensity and suicidality. Participants who endorsed higher rates of eveningness showed higher risk taking propensity ($p < .05$) and suicidality ($p < .001$). Eveningness was a significant predictor of suicidal ideation ($p = .05$). Explanations of these findings and the subsequent implications will be discussed at length.

Keywords: circadian preference, sleep chronotype, suicidality, emotion regulation

The transition to adulthood is a growing period characterized by changes in physical, mental, and emotional health. During the transition to adulthood young adults exhibit superior health relative to other periods; however, it is also a critical period for the development of healthy behavioral habits, and can be a defining period for development of later negative health outcomes. Susceptibility to psychopathological illnesses is highest during this period (Dahl, 1996; Cicchetti, Rogosch, & Toth, 1994). Recent research also highlights increasing rates of childhood obesity occurring during this critical period (Ebbeling, Pawlak, & Ludwig, 2002; Kosti & Panagiotakos, 2006).

During this period circadian rhythms are particularly vulnerable to various external and internal influences which significantly affect health during later life. Those who are transitioning to adulthood require more sleep on average than older adults (Carskadon, 1990). Sleep deprivation during this phase can lead to lowered academic performance (Gruber et al., 2010), decreased performance in memory tasks (Lim & Dinges, 2010), decreased ability to regulate mood (Talbot, McGlinchey, Kaplan, Dahl, & Harvey, 2010), and poor dietary habits (Nedeltcheva, Kilkus, Imperial, Schoeller, & Penev, 2010).

Evening Circadian Tendency The current study aims to examine a related, but often overlooked, sleep construct that may also contribute to negative health outcomes—that is, an evening circadian tendency. Individuals with an evening

circadian tendency (E-types) are phase advanced, sleeping long into the day and becoming most active at night. In comparison, individuals with a morning circadian tendency (M-Types), prefer to wake early in the morning and find it challenging to remain awake past their ordinary bedtime (Webb & Bonnet, 1978). Like most people in the transition to adulthood period, E-types operate on a daily fixed schedule and must rise early in the morning; this often results in sleep deprivation. School usually starts early in the morning for most teens which further exacerbates the problem (Wolfson, Spaulding, Dandrow, & Baroni, 2007). As sleep phase is delayed the biological clock shifts as well. This shift has an effect on biological functions that rely on a normal sleep phase such as metabolism and cognitive functioning in the frontal lobe (Carskadon, Acebo, & Jenni, 2004).

Carskadon and colleagues (1993) hypothesized that there may be a link between biological mediation of phase delay and puberty. The view is theoretically plausible because melatonin, the neurochemical compound that plays a vital role in the regulation of the sleep-wake cycle, has been found to induce the production of gonadotrophin-inhibitory hormone (GnIH), a hormone that inhibits the production of gonadotropine releasing hormone (GnRH) (Ubuka, Bentley, Ukena, Wingfield, & Tsutsui, 2005). GnRH is broadly responsible for the development of the gonads starting at puberty. Melatonin has been found to have an inverse relationship with GnRH secretion over the span of pubertal development (Tsutsui, 2005). This means that we could naturally expect to see huge links in timing of sleep phase during pubertal development just as Carskadon (1993) suggested even though this sort of research is in its infancy.

Delayed sleep phases are already known to affect social, behavioral, and emotional function in teens. It leads to poor academic performance (Randler & Frech, 2009; Wolfson & Carskadon, 2003), increased impulsivity (Adan, Natalie, Caci, & Prat, 2010), poor emotional regulation (Gau, Shang, Merikangas, Chiu, Soong, & Chang, 2007) and poor eating habits (Nedeltcheva et al., 2010). Delayed sleep phases have also been linked with various psychopathological illnesses, including depression (Campos-Hirata, Lima, de Bruin, Nobrega, Wenceslau, and de Bruin, 2007; Randler, 2011); anxiety disorder (Gau et al., 2007), suicidality (Gau et al., 2007) and ADHD (Caci, Bouchez, & Bayle, 2009). Furthermore, past research has shown that delayed sleep phases place many adolescents at moderate to high risk for obesity, heart disease, cancer and immune system malfunction in adulthood (Ayas, White, Manson et al., 2003; Kanterman & Ronnenberg, 2009).

Study Aims The current study will investigate the extent to which E-types and M-types differ in response to items measuring suicidality and risk propensity and the extent to which being an E-type can predict suicidality and risk propensity. There is good reason to believe that E-types and M-types will differ in suicidality and risk propensity given past research that has demonstrated that E-types and M-types differ biologically (e.g. in terms of their physiological rhythms), and emotionally (e.g. in terms of emotion regulation). E-types and M-types also differ on the endogeneity of their sleep phases such that the cyclic rhythms of physiological variables such as alertness, core body temperature, heart rate, and hormone secretion all vary as a function of sleep chronotype (Baehr et al. 2000; Monk et al., 1997; Bailey & Heitkemper, 1991; Giannotti, Cortesi, Sebastiani, and Ottaviano, 2002). Behavioral and performance rhythms have also been noted to differ across sleep chronotypes (Carrier & Monk, 2000). In a study to determine the relationship between circadian preferences, regularity of sleep patterns, sleep problems, and daytime behavior Giannotti and colleagues (2002) assessed a nationally representative sample of Italian adolescents ranging from 14.1-18.6 years of age. One of the studies' findings was that E- and M-types differ significantly in the reporting of emotional problems such that having an evening circadian preference significantly increased the likelihood of reporting emotional difficulties. For these reasons we compare a large nationally representative sample of E-types and E-types on indices of suicidality and risk propensity. We evaluate data from the National Longitudinal Study of Adolescent Health (NLSAH) which assessed a nationally representative sample of adolescent students in America across 14 years into adulthood.

METHOD

Participants and Recruitment The current study used data from the National Longitudinal Study on Adolescent Health (ADD Health). ADD Health is a national longitudinal survey of school-aged students in grades 7-12 that ran from 1994 to 2008. Four separate waves of data were collected over a span of 15 years. The survey objective was to investigate how health-related behaviors in adolescence affected behaviors and wellbeing in later adulthood.

Survey procedures (Resnick et al., 1997) were reviewed and approved by the Institutional Review Board at the University of North Carolina. The first wave of data (1994-1995) focused on factors that may influence adolescent health and risk factors; it included more than 90,000 students and over 100

school administrators who completed surveys administrated in school. A subset of 20,745 adolescents and 17,670 parents completed in-home surveys. The second wave of data (1996) included 14,738 of the adolescents from the first wave who had yet to graduate from high school, including those who dropped out. The third and fourth waves of data (2001-2008) focused on how experiences during the transition to adulthood (i.e. ages 18-21) are related to decisions, health, and behavior in later adulthood each included more than 15,000 of the adolescents from the first wave.

OUTCOME MEASURES

Risk Propensity Risk propensity can loosely be defined as the degree to which a person is willing to take a chance or react with behavioral impulsivity with either regard or complete disregard for the consequences of that action. The construct of risk propensity was measured using two survey items. Participants first responded to the statement, "I like to take risks," using a 5-point Likert scale. Then, they responded to the second statement, "When making a decision I tend to go with my 'gut feeling' without regard for the consequences that may follow my decision", which was also measured on the same Likert scale ranging from 1(strongly agree) to 5 (strongly disagree). Each of these two variables was included in Wave III and Wave IV surveys so we would be able to identify trends in these data.

Suicidality Suicidality is defined as the act of engaging in suicidal ideation of any sort. The construct of suicidality was assessed using two survey items. Participants first responded to the question, "In the past 12 months have you thought about committing suicide?" Answers were recorded on a nominal scale and coded 0 (no) and 1 (yes). Participants then responded to the second question, "In the past 12 months how many times have you actually attempted to commit suicide?" Answers were recorded on an ordinal scale ranging from 1 (once) to 5(five or more times). Each of these items was included in Wave III and Wave IV surveys.

PROCEDURE

Defining E-types and M-types Based on previous research, we utilized participants' sleep-wake times as proxies for chronotype (Mullin, Harvey, and Hinshaw, 2011; Eidelman et al., 2010). We derived an estimate of chronotype from two Wave III (2002) survey items: (1) 'On days you don't have to get up at a certain time, what time do you usually get up?' and (2) 'On those days, what time do you usually go to sleep the night or day before?'   Based   on   previous   research (Roenneberg, 2003; Benoit et al., 1981) we selected the midpoint between sleep onset and wake up time (midsleep) as an estimate of chronotype. Midsleep has also been reported as the best phase anchor point for melatonin onset— the period of time during the day that the body prepares itself for sleep most notably through the secretion of higher levels of the neurohormone (Terman et al., 2001). We then performed a tertile split on our sample based on midsleep such that we labeled the lower third as E-types —those with delayed sleep-wake times, and the upper third as M-types —those who have early sleep-wake times. We did not include the middle third of the sample population in an attempt to avoid drawing arbitrary lines right down the middle between E- and M-type sleep-wake times. We used the tertile split to get rid of the gray area in between E-types and M-types —those participants who fell into some middle ground between E- and M-type sleep-wake times. As a result of excluding this group of participants, we can be more confident that those participants who are included as E- and M-types are accurately placed in each of the two groups, thus, strengthening our comparisons between the sleep chronotypes.

Statistical Analysis The primary aim of the current study is to see if being an E-type in Wave III could predict suicidal ideation and risk propensity in Wave IV. Ordinal and binary logistic regressions were used to evaluate the effects that being an E-type had on suicidality and risk propensity based on the aforementioned survey items utilized in this study. To protect the statistical validity of our regression analysis we included participant responses from Wave III as predictors for answers to those same questions in Wave IV.   We appreciated the fact that participant responses from the past may be significant predictors to their responses later. Simply put, we were in search of how strong the predictive influence of being an E-type was on suicidality and risk propensity when accounting for participant responses to Wave III items. Chi-square tests were also used to evaluate the differences between E-types and M-types on a number of factors such as, race, and gender; t-tests were also used to compare difference in age between the two groups. An alpha level of 0.05 was used for all analyses.

Table 1. Comparative Demographics			
	E-Type ($n = 1022$)	**M-Type** ($n = 1112$)	**Test Statistic**
Female	507	698	$x^2 = 37.53***$
Male	515	414	
Mean Age (SD)	20.04 (.83)	20.13 (.82)	$t = -2.46**$
Age Range	18-21	18-21	
Ethnicity			
Asian	44	45	$x^2 = 0.08$
American Indian	40	51	$x^2 = 0.59$
Caucasian	705	741	$x^2 = 1.29$
Hispanic	88	124	$x^2 = 3.83*$
African American	238	281	$x^2 = 1.16$
Other	52	58	$x^2 = 0.02$
$p < 0.05*$, $p < 0.01**$, $p < 0.001***$			

Table 2. Chi-Square Analysis of Survey Responses between M-Type and E-Type		
Wave III Item	x^2	**Significance**
I like to take risks	28.38	.00**
When making a decision I usually go with my 'gut feeling' without much thought about the consequences of that decision.	8.09	.08
During the past 12 months, have you ever seriously thought about committing suicide?	17.38	.00**
During the past 12 months, how many times have you attempted suicide?	2.79	.72
Wave IV Item	x^2	**Significance**
I like to take risks	16.03	.00**
When making a decision I usually go with my 'gut feeling' without much thought about the consequences of that decision.	9.03	.06
During the past 12 months, have you ever seriously thought about committing suicide?	8.03	.00**
During the past 12 months, how many times have you attempted suicide?	7.30	.12
$p < 0.05*$, $p < 0.01**$, $p < 0.001***$		

RESULTS

Participants We utilized subsets of the samples from Wave III and Wave IV. Our sample only included participants who ranged in age from 18-21 in Wave III. Participants who completed both Wave III and Wave IV were included in the sample. Thus, the final sample size was 2,134. Table 1 shows detailed demographics of the subsample used in our study. We observed a statistically significant difference in age between the two groups (t=-2.46, p<.01) such that the average age of the E-types was 20.04 years (SD=.83), and the average age of the M-types was 20.13 (SD=.83). However, even though this result was statistically significant we are fairly certain that it is too small to be of any practical significance.

Group Differences in Sex and Race Table 1 also shows results from chi-square analyses for observed differences between E-types and M-types on indices of sex and race. We observed that there were significantly fewer female E-types, and fewer male M-types than expected; the data also showed that there were more female M-types and more male E-types than expected (x^2 = 37.53, p<.001). This suggests that there is a strong association between being female and being an M-type, and being male and being an E-type. Table 1 also reports statistical analysis on differences between E- and M-types in various ethnic groups. We found that there are significantly fewer E-types that were of Hispanic origin than expected, and more M-types that were of Hispanic origin than expected (x^2 = 3.83, p =.05). In all other chi-

Table 3. Ordinal Regression Analysis of E-Type on Wave IV Responses					
Wave IV Item	**R²**	**Predictors**	**Odds Ratio**	**Significance**	**CI**
I like to take risks	.21	E-Type	.86	.08	-.32 -.02
		Wave III item	2.38	.00***	.78 - .95
When making a decision I usually go with my 'gut feeling' without much thought about the consequences of that decision.	.11	E-Type	.96	.65	-.21 - .13
		Wave III item	1.74	.00***	.47 - .63
During the past 12 months, how many times have you attempted suicide?		E-Type	1.03	.71	-1.50 - 1.58
	.16	Wave III item	3.61	.00***	.58 – 1.58
p < 0.05*, p<0.01**, p < 0.001***					

Table 4. Binary Logistic Regression of E-Type on Wave IV Suicidal Ideation item					
Predictors	**B**	**Std. Error**	**Odds Ratio**	**Significance**	**CI**
E-Type	.38	.19	1.47	.05*	.99 – 2.16
During the past 12 months, have you seriously thought about committing suicide?	-1.78	.21	.169	.00***	.11 - .25
Constant	-1.35	.23	.25	.00***	
p < 0.05*, p<0.01**, p < 0.001***					

square analyses there were no differences between E-types and M-types regarding ethnicity.

Group Differences in Suicidality and Risk Propensity Table 2 shows chi-square analysis for observed differences between E-types and M-types on indices of suicidality and risk propensity based on self-reported responses to the items used to measure each construct. We found that significantly more E-types seriously thought about committing suicide during the 12 month period before they participated Wave III (x^2 = 17.38, p<.01) and in Wave IV (x^2 = 8.03, p<.01) during later adulthood compared to M-types . However, the data showed no significant difference in the amount of times E-types actually attempted suicide during the 12 month period before they participated in Wave III (x^2 = 2.79, p =.72) and Wave IV(x^2 = 7.30, p =.12). Regarding risk propensity, the data showed marginally significant differences between E-types and M-types, such that, E-types reported being more likely to make a decision without regard for the consequences in later adulthood compared to M-types in Wave III (x^2 = 8.09, p =.08) and in Wave IV (x^2 = 9.03, p =.06). The data also showed that E-types are more likely to report that they enjoy engaging in risks in later adulthood compared to M-types at Wave III (x^2 = 28.38, p<.01) and at Wave IV (x^2 = 8.03, p<.01).

Long-term Relationship: E-type, Suicidality, and Risk Propensity Table 3 shows the ordinal regressions for all three of our ordinal survey items in Wave IV. Identical survey items that participants responded to in Wave III were also included in the model so that the true predictive power of being an E-type during Wave III would be displayed in the each ordinal regression models. The data showed that being an E-type during Wave III was a marginally significant predictor (OR=.86, CI= -.32 - .02, p =.08) of responses to the Wave IV statement, "I like to take risks," over and above their response to the same Wave III survey item that was administered at an earlier time point which was a significant predictor (OR=2.38, CI= .78 - .95, p <.01). Being an E-type during Wave III was not a significant predictor of responses to the Wave IV statement, "When making a decision I usually go with my 'gut feeling' without much thought about the consequences of that decision," (OR= .96, CI= -.21 - .13, p =.65). The identical item in Wave III was a significant predictor however (OR=1.74, CI= .47 - .63, p <.01). Regression analysis also showed that being an E-type during Wave III was not a significant predictor of responses to the Wave IV question, "During the past 12 months, how many times have you attempted suicide," (OR=1.03, CI= -1.50 – 1.58, p =.71). However, the responses to the identical Wave III statement was found to be a

significant predictor of answers to the same question in Wave IV (OR =3.61, CI =.58 – 1.58, p <.01).

We ran a binary logistic regression for the suicidal ideation item shown in Table 4. Analysis shows that an E-type during Wave III was indeed a statistically significant predictor of the responses to the Wave IV suicidal ideation item over and above the predictive power of the Wave III questions (OR =3.61, CI = .99 – 2.16, p =.05). Participants responded to the statement, "During the past 12 months, how many times have you thought about committing suicide." We found that the corresponding Wave III item was also a significant predictor (OR =.169, CI = .11 - .25, p <.01).

DISCUSSION

In line with our first aim, we found significant differences between E-types and M-types in responses to Wave III and in Wave IV items such that E-types tend to report taking risks and report seriously considering suicide more than M-types. However, we note that we cannot draw on any causal explanations for the significance we found when comparing E- and M-types through chi-square analyses. In regards to our second aim, being an E-type in Wave III was found to be a significant predictor of suicidal ideation in Wave IV over and above the Wave III response to the same question as recorded four years prior. The data suggests an important association between being an E-type during the transition to adulthood, suicidality, and risk propensity.

All of these results suggest that circadian rhythms during the transition to adulthood should be monitored closely, and attempts should be made to align their sleep-wake time with their early morning schedules. Although sleep-wake preference seems to be endogenous, it is malleable to an extent that has been found to be advantageous to those who are E-types (Skene & Arendt, 2006). Gross fixed daily schedules are not conducive to students who are considered E-types; the evidence suggests that this leads to sleep deprivation in E-types and puts them at a disadvantage academically, socially, and emotionally (Cassoff, & Monson, 2010; Talbot, McGlinchey, Kaplan, Dahl, and Harvey, 2010). This warrants some attention, given the evidence that being an E-type leads to sleep deprivation which further leads to compromised adult health outcomes such as immune system deregulation; hypertension; obesity; mood disorders; and cognitive decline (Durmer & Dinges, 2005; Gottlieb et al., 2006; Colton & Altevogt, 2006; Lange, Dimitrov, and Born, 2010; Benedetti & Colombo, 2011).

Future studies should investigate the influence that circadian preference during earlier developmental years has on long-term health as well as collect for extensive data on outcome measure. Due to the fact that we analyzed data from a public research database we did not have control over what survey items were included in the study and when those survey items were administered. Despite the limitations, the current study reveals a clear influence of circadian preference on behavior.

Some of the fetal programming literature on animals suggests that prenatal stress can induce a circadian phase advance of locomotor activity as a result of modifications made to the hypothalamic-pituitary adrenal axis (Maccari, Darnaudery, Morely-Fletcher, Zuena, Cinque, and Van Reeth, 2003; Van Reeth, Dugovic, and Koehl, Weibel, and Maccari, 1999). This is highly relevant because it hints at the fact that circadian rhythms of infants may be influenced by prenatal stress, leaving some infants disadvantaged in regards to their circadian preference before they are introduced to the world. Thus, it may be insightful to track a nationally representative sample of students from prenatal to adulthood years while collecting data on in utero stress hormone exposure, indices of sleep chronotype, and a myriad of other health factors. We also hope that future research will gather more extensive outcome measures, as it is beneficial to have multiple facets of a single construct represented in varying survey items.

In conclusion, these findings illustrate that sleep-wake patterns during the transitional stage to adulthood can importantly shape a person's future mental health. Further studying this relationship will prove important for devising policy interventions to help combat these negative effects.

References

Adan, A., Natale, V., Caci, H., & Prat, G. (2010). Relationship between circadian typology and functional and dysfunctional impulsivity. *Chronobiology International* , (27), 606-619.

Ayas, N. T., White, D. P., & Manson, J. E. (2003). A prospective study of sleep duration and coronary heart disease in women. *Archives of Internal Medicine*, (163), 205-209.

Baehr, E. K., Revelle, W., & Eastman, C. J. (2000). Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness-eveningness. *Journal of Sleep Research*, *9*, 117-127.

Bailey, S., & Heitkemper, M. (1991). Morning-evening and early morning salivary cortisol levels. *Biological Psychology*, *32*, 181-192.

Benedetti, F., & Colombo, C. (2011). Sleep deprivation in mood disorders.*Neuropsychobiology*, *64*(3), 141-151.

Benoit, O., Foret, J., Merle, B., & Reinberg, A. (1981). Circadian rhythms (temperature, heart rate, vigilance, mood) of short and long sleepers: effects of sleep deprivation. *Chronobiologia, 8*(4), 341-350.

Borisenkov, M. F., Perminova, E. V., & Kosova, A. L. (2010). Chronotype, sleep length, and school achievement of 11- to 23-year-old students in northern European Russia.*Chronobiology International* , *27*, 1259-1270.

Caci, H., Bouchez, J., & Bayle, F. J. (2009). Inattentive symptoms of ADHD are related to evening orientation. *Journal of Attention Disorders* , (13), 36-41.

Campos-Hirata, F., Lima, M. C., De Bruin, V. M., Nobrega, P. R., Wenceslau, G. P., & De Bruin, P. F. (2007). Depression in medical school: The influence of morning-eveningness. .*Chronobiology International* , *24*, 939-946.

Carrier, J., & Monk, T. (2000). Circadian rhythms of performance: new trends. *Chronobiology International*, *17*, 719-732.

Carskadon, M. A., & Acebo, C. (1993). A self-administered rating scale for pubertal development. *Journal of Adolescent Health*, *14*, 190-195.

Carskadon, M. A., Acebo, C., & Jenni, O. G. (2004). Regulation of adolescent sleep: Implications for behavior. *Annals of the New York Academy of Science* , (1021), 276-291.

Carskadon, M. A., Mindell, J. A., & Drake, C. (2006). Sleep in America Poll. *The National Sleep Foundation*.

Carskadon, M. A. (1990). Patterns of sleep and sleepiness in adolescents. *Pediatrician, 17*(1), 5-12.

Carskadon, M. A. (2002). Adolescent sleep patterns: Biological, social, and psychological influences . In *Factors influencing sleep patterns in adolescents*. (16th ed.). (pp. 4-26). New York, NY: Cambridge University Press.

Cicchetti, D., Rogosch, F. A., & Toth, S. L. (1994). *A developmental psychopathology perspective on depression in children and adolescents*. (1st ed.). New York, NY: Plenum Press.

Colten, H. R., & Altevogt, B. M. (Eds.). (2006). *Sleep Disorders and Sleep Deprivation: An unmet Public Health Problem*. Washington, D.C.: National Academies Press.

Colten, H. R., & Altevolt, B. M. (Eds.). (2006). *Sleep disorders and sleep deprivation: An unmet public health problem*. (1st ed.). Washington, D.C.: National Academies Press.

Dahl, R. E. (2002). The regulation of sleep-arousal, affect, and attention in adolescence: Some questions and speculations. In M. A. Carskadon (Ed.). *Adolescent sleep patterns: biological, social and psychological influences*. (1st ed.).

(pp. 269-284). Cambridge, Massachusetts : Cambridge University Press.

Dahl, R. E. (1996). The regulation of sleep and arousal :Development and psychopathology.*Development and Psychopathology*, *8*, 3-27.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation.*Seminars in Neurology*, *25*(1), 117-129.

Ebbeling, C. B., Pawlak, D. B., & Ludwig, D. S. (2002). Childhood obesity: public-health crisis, common sense cure. *The Lancet*, (360), 473-482.

Eidelman, P., Talbot, L. S., Gruber, J., & Harvey, A. G. (2010). Sleep, illness course, and concurrent symptoms in inter-episode bipolar disorder. *Journal of Behavioral Therapy and Experimental Psychiatry, 41*(2), 145-149.

Gau, S. S., Shang, C. Y., Merikangas, K. R., Chiu, Y. N., Soong, W. T., & Cheng, A. T. (2007). Association between morningness-eveningness and behavioral/emotional problems among adolescents. *Journal of Biological Rhythms*, *22*, 268-274.

Giannotti, F., Cortesi, F., Sebastiani, T., & Ottaviano, S. (2002). Circadian preference, sleep and daytime behaviour in adolescence. *Journal of Sleep Research*, *11*, 191-199.

Gottlieb, D. J., Redline, S., Nieto, F. J., Baldwin, C. M., Newman, A. B., Resnick, H. E., & Punjabi, N. M. (2006). Association of usual sleep duration with hypertension: the sleep heart health study. *Sleep*, *29*(8), 1009-1014.

Gray, E. K., & Watson, D. (2002). General and specific traits of personality and their relation to sleep and academic performance. *Journal of Personality*, (70), 177-206.

Gruber, R., Laviolette, R., Deluca, P., Monson, E., Cornish, K., & Carrier, J. (2010). Short sleep duration is associated with poor performance on IQ measures in healthy school-age children. *Sleep Medicine, 11*(3), 289-294.

Gruber, R., Wiebe, S. T., Wells, S. A., Cassoff, J., & Monson, E. (2010). Sleep and academic success: mechanisms, empirical evidence, and interventional strategies. *Adolescent Medicine State Art Review, 21*(3), 522-541.

Kantermann, T., & Roenneberg, T. (2009). light-at-night a health risk or a health predictor? .*Chronobiology International*, (26), 1069-1074.

Kosti, R. I., & Panagiotakos, D. B. (2006). The epidemic of obesity in children and adolescents in the world. *Central European Journal of Public Health* , *14*, 151-159.

Lange, T., Dimitrov, S., & Born, J. (2010). Effects of sleep and circadian rhythm on the human immune system. *Annals of the New York Academy of Sciences*, *1193*, 48-59.

Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin, 136*(3), 375-389.

Maccari, S., Darnaudery, M., Morley-Fletcher, S., Zuena, A. R., Cinque, C., & Van Reeth, O. (2003). Prenatal stress and long-term consequences: implications of glucocorticoid hormones. *Neuroscience and biobehavioral reviews*, *27*(1-2), 119-127.

Monk, T. H., Busse, D. J., Reynorld III, C. F., Berga, S. L., Jarre-Bergley, A. E., & Kupfer, D. J. (1997). Circadian rhythms in human performance mood under constant conditions.*Journal of Sleep Research*, *6*, 9-18.

Mullin, B. C., Harvey, A. G., & Hinshaw, S. P. (2011). A preliminary study of sleep in adolescents with bipolar disorder, ADHD, and non-patient controls. *Bipolar Disorder, 13*(4), 425-432.

Nedeltcheva, A. V., Kilkus, J. M., Imperial, J., Schoeller, D. A., & Penev, P. D. (2010). Insufficient sleep undermines dietary efforts to reduce adiposity. *Annals of Internal Medicine, 153*(7), 435-441.

Randler, C., & Frech, D. (2006). Correlation between morningness - eveningness and final school leaving exams. *Biological Rhythm Research* , (37), 233-239.

Randler, C., & Frech, D. (2009). Young people's time-of-day preferences affect their school performance. *Journal of Youth Studies* , *12*, 653-667.

Randler, C. (2011). Association between morningness-eveningness and mental and physical health in adolescents. *Psychological Health Medicine, 16*(1), 29-38.

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., . . . Udry, J. R. (1997). Protecting adolescents from harm. Findings from the National Longitudinal Study on Adolescent Health. *JAMA, 278*(10), 823-832.

Roenneberg, T., Wirz-Justice, A., & Merrow, M. (2003). Life between clocks: daily temporal patterns of human chronotypes. *Journal of Biological Rhythms, 18*(1), 80-90.

Roenneberg, T., Kuehnle, T., & Pramstaller, P. P. (2004). marker for the end of adolescence.*Current Biology*, (14), R1038.

Skene, D. J., & Arendt, J. (2006). Human circadian rhythms: physiological and therapeutic relevance of light and melatonin. *Annals of Clinical Biochemistry*, *43*, 344-353.

Talbot, L. S., McGlinchey, E. L., Kaplan, K. A., Dahl, R. E., & Harvey, A. G. (2010). Sleep deprivation in adolescents and adults: changes in affect. *Emotion, 10*(6), 831-841.

Terman, J. S., Terman, M., Lo, E. S., & Cooper, T. B. (2001). Circadian time of morning light administration and therapeutic response in winter depression. *Archives of General Psychiatry, 58*(1), 69-75.

Tsutsui, K. (2005). Gonadotrophin-Inhibitory Hormone (GnIH). *Japanese Journal of Reproductive Endocrinology*, 10, 49-53.

Ubuka, T., Bentley, G. E., Ukena, K., Wingfield, J. C., & Tsutui, K. (2005). Melatonin induces the expression of gonadotrophin-inhibitory hormone in the avian brain. *Proceedings of the National Academy of Sciences*, 102(8), 3052-3057.

Van Reeth, O., Dugovic, C., Koehl, M., Weibel, L., & Maccari, S. (1999). Hormonal and behavioral effects of prenatal stress: focus on circadian rhythms and sleep. *Journal of Sleep-wake Research*, 39-50.

Webb, W. B., & Bonnet, M. H. (1978). The sleep of 'morning' and 'evening' types. *Biological Psychology, 7*(1-2), 29-35.

Whittmann, M., Dinich, J., Merrow, M., & Roenneberg, T. (2006). Social jetlag: Misalignment of biological and social time. *Chronobiology International*, (23), 497-509.

Wolfson, A. R., & Carskadon, M. A. (2003). Understanding adolescents' sleep patterns and school performance: a critical appraisal. *Sleep Medicine Review, 7*(6), 491-506.

Wolfson, A. R., Spaulding, N. L., Dandrow, C., & Baroni, E. M. (2007). Middle school start times: the importance of a good night's sleep for young adolescents. *Behavioral Sleep Medicine, 5*(3), 194-209.

Testing Efficacy of Sensorimotor Enrichment in Ameliorating Symptoms of Rett Syndrome in a Mouse Model

Mahi Worku Megra

Georgetown University

Pamela Diener, PhD, MS, OT/L, Principal Investigator

Georgetown University Medical Center and Marymount University

Rett Syndrome (RTT) is a genetic, neurodevelopmental disorder that primarily affects girls (Hagberg, Aicardi, Dias, & Ramos, 1983). A mutation of the *Mecp2* gene has been implicated as a cause of RTT (Amir et al., 1999). Between 6-18 months of age, individuals with RTT experience deterioration in various cognitive, motor and social skills such as walking, speaking, and pointing (Naidu, 1997). Recent research of motor recovery hints that sensorimotor enrichment helps to improve the motor skills of *Mecp2* null mice (Kondo et al., 2008). Our study was performed using wild type mice and mutated RTT mice. Some of the mutated mice were exposed to daily enrichment. Our study not only used novel enriched-environments targeted at particular motor skills, but also crafted two motor tests/criteria for scoring that directly assess specific component motor skills necessary for functional activity. Our findings show that sensorimotor enrichment may ameliorate the motor deficiencies seen in this RTT mouse model. In addition, the results suggest that discontinuing enrichment may lead to deterioration of motor skills and reinstitution of enrichment may support motor skill recovery as assessed by these two tests. These results suggest that enrichment might be a viable treatment method for enhancing motor skill performance in girls with RTT.

Rett syndrome (RTT) is a neurodevelopmental disorder that affects 1 in 10,000 females (Amir et al., 1999). The clinical manifestations of RTT include the deterioration of acquired cognitive, motor and language skills between the ages of 6 and 18 months and the loss of purposeful hand use (Naidu, 1997). Initially, it was suggested that RTT might be an X-linked genetic disorder that mainly affects girls (Hagberg et al., 1983). A familial analysis of monozygotic and dizygotic twins provided compelling evidence that the disorder was in fact associated with the X-chromosome (Percy, 1992). However, it was not until 1999 that the specific locus for a mutation in the X-chromosome that might potentially dictate the pathogenesis of RTT was reported (Amir et al., 1999). Amir et al. (1999), identified mutations in the gene for methyl CpG-binding protein 2 (MeCP2) at the locus Xq28 as the cause of most RTT cases. *Mecp2* is a transcription regulator that targets many genes such as the *bdnf* gene (Chen et al., 2003).

The *bdnf* gene produces the protein brain-derived neurotrophic factor (BDNF), which helps to develop and maintain dendritic connections in immature and mature nervous systems (Chen et al., 2003). The regulation of BDNF is complex, involving the phosphorylation of MeCP2. Mutations of MeCP2 inhibit it from being phosphorylated, thereby preventing the eventual production of the BDNF protein (Sun & Wu, 2006). Due to the mutation in the *Mecp2* gene in people with RTT, low levels of brain BDNF levels have been reported (Moretti & Zoghbi, 2006).

There are multiple invasive ways to increase BDNF levels in the CNS. One noninvasive method seems to arise when environmental enrichment is present. Effects of enrichment have shown to be beneficial in the recovery of lost motor skills in multiple rat and mouse models of disease/injury (Diener, 2002;

Klinstova et al., 1998: Berrocal et al., 2007; Kondo, et al. 2008). Currently, there is no standard intervention for the treatment of RTT. However, environmental enrichment has been suggested as an option to ameliorate the symptoms experienced by *Mecp2* knockout mice via increasing BDNF levels (Kondo, et al. 2008). Sensorimotor or environmental enrichment, in this specific context, means housing the mice in an environment that is stimulating and includes play items with different textures. The present study tested whether the symptoms experienced by *Mecp2* knockout mice could be minimized or eradicated after multiple days of exposure to varied play environments, each providing sensorimotor enrichment. An additional purpose of the study is to create and evaluate two novel testing batteries to determine how effectively they can assess the performance of mice with RTT-like symptoms.

In previous studies using mouse models of neurologic diseases (Rett Syndrome*,* Spinal Cord Injury, Huntington's disease), assessment of motor control was conducted using various tests such as the accelerating rotarod (Kondo et. al, 2008; Tabuse, Yaguchi, Ohta, Kawase & Toda, 2010; Berrocal et al., 2007). The rotarod is an automated cylinder that accelerates at a fixed rate. For testing, the mouse is placed on the surface and time spent on the rod is recorded in order to examine how long the animals remain on the surface. We opted against using the rotarod because it is designed to test generalized motor abilities such as coordination and balance (Shiotsuki et al., 2010) rather than isolating and challenging the recovery of specific components of skilled function as has been previously reported in another model of immature nervous system injury (Diener and Bregman, 1998). In addition, testing using the rotarod shows that sometimes mice remain on the accelerating rod while gripping it tightly, rotating around with the rod, without actually moving their forelimbs and hindlimbs to remain balanced upright on the rod. Since the assessment scale for this battery is mainly time spent on the rod rather than measuring motor skills used to remain on the rod, we believe that the rotarod may not be effective in measuring improvements in/recovery of motor skills in mouse models of Rett Syndrome. Hence, we assessed performance using modified alternate batteries i.e., the Grid Walk and Grid Hang (Carlson, 2012) in which we developed a scoring criteria to meticulously analyze the movement and patterns of movement.

METHOD

Animals Male strain B6.128S-MeCP2[tmlHzo/J] mice (Jackson Labs, Maine) were obtained after weaning. There were three groups of mice: *Mecp2* mutant mice exposed to the sensorimotor enrichment environment (SMEE) (E-group; n=5), *Mecp2* mutant mice that were not exposed to SMEE (N-group; n=4), and wild type mice, which were also not exposed to SMEE (W-group; n=6). The mice were housed in the Georgetown University Division of Comparative Medicine (DCM) and had access to unlimited food and water. Two mice in the E-group died during the study of causes seemingly unrelated to Rett syndrome. The Georgetown University Animal Care and Use Committee approved all animal protocols.

Study Design Baseline measurements were obtained four days after the mice arrived at the DCM facility. SMEE started three days after baseline testing. Mice were placed into the enriched environments five days a week, for four hours a day during the intervention period of the study. The study design includes an enrichment period for the first 7 months (March 19-October 1) followed by three month period when enrichment is discontinued (October 2 – Dec 28) and a final third period when enrichment is reinstituted (December 29-February 15). Behavioral testing took place every 4 weeks throughout the entire study, except in February when the final testing was scheduled 6 weeks after reinstitution of enrichment.

Environmental Enrichment Different tactile cues and motivating toys that elicited play and digging behaviors along with varied textured terrain and climbing equipment to encourage exploration were used within the 10 different sensorimotor enriched environments. Each environment was designed to spare the deterioration of or promote the recovery of different aspects of motor behavior typically lost in RTT mutant mice. Each mouse was exposed to each of the environments for a 4 hour time period for 5 days /week. This typically increased novelty and encouraged the mouse to explore the environment throughout the time that each was within, rather than exploring and then falling asleep in the corner of a cage. The animals were constantly monitored to ensure that they were engaged actively in the novel environment and daily notes of activity in each SMEE were recorded.

Behavioral Testing and Measures All testing protocols were videotaped with Canon Vixia HG20 camcorder with a shutter speed of 1/1000 sec and later analyzed using MacBook Pro laptops QuickTime player through frame-by-frame analysis. This study looks at two specific kinds of behavioral testing. To focus more specifically on testing particular aspects of motor skills and posture/balance, we chose to modify tests that would give a clear interpretation of recovered motor skills (or recovered components of motor skills).

Based on previous work, the tail hang test was used to validate *Mecp2* mutation through stereotypical forelimb movements (Young & Zoghbi, 2004). During this test, the researcher suspended each mouse by his tail to observe for a forelimb clasping. If forelimb clasping was identified, the animal received a score of 2. If clasping was absent, a score of 1 was given.

We modified two tests of motor skills that are widely used to study the rodent's ability to maintain grip against gravity, *Grid Hang* and achieve precise foot placement during ambulation, *Grid Walk*.

A) Grid Hang The behavioral test of grid hang was modified from Tillerson and Miller's study (2003) assessing mice affected with Parkinson's Disease. The grid hang test measures the core strength, coordination and skilled limb movements of mice. Our testing apparatus was constructed from a grid rack (37 cm x 50 cm; each opening of grid is 1 cm x 1 cm). The test begins by placing the mouse on the grid. Once the mouse grips the grid, the grid is inverted over a large bin filled with multiple layers of bedding materials. The mouse is required to hold its body on the grid while hanging upside down. The grid size was chosen because it was big enough to allow the animals to grasp the bars without difficulty but small enough for the mice to transition between the bars with ease and to move around on the grid, if desired. To compare the performance of the three groups of mice, we measured the time (in seconds) that the animals gripped the grid bars to maintain their body position against gravity. Testing was terminated after 60 seconds or when they fell onto the cushioned surface below them, whichever occurred first. This test appears to be a good measurement of the core strength of the animal. Testing was repeated three times to maximize the number of results collected and to minimize a change being due to an error.

B) Grid Walk This protocol was also adapted from a study that used the testing for other mouse models (Cummings, Engesser-Cesar, & Anderson, 2003; Merkler, 2001; Pajoohesh-Ganji, Byrnes, Fatemi, & Faden, 2010). The behavioral testing apparatus was constructed like the rungs of a ladder using six-inch long wooden dowels spaced 1 cm apart on a runway that was approximately 1 meter long. The object was for each mouse to walk from one end of the grid to the other. The spacing between the dowels was determined in random order (A coin was flipped to determine whether the dowel remained or was removed) so that the mouse could not predict what would happen next on each testing days. Testing took place in horizontal, ascending and descending planes. To test equal loading of forelimbs and hindlimbs, the grid was placed along a horizontal surface. The two other testing scenarios involved a grid declined or inclined 55 degrees. Hence, when the animal walked up the ladder-like structure, the hindlimbs were loaded. In contrast, when the animal descended the grid, the forelimbs were loaded. This test was constructed primarily to test for coordination and proprioception deficits.

This study used a unique measurement scale to evaluate the method of grip and non-grip on the grid walk within each of three planes. Non-grips were operationalized as a stutter (step into the air/space between rods), repeated attempts for forward progression to the same rod, flexing of fingers without a dowel (gripping the air), and a full or partial limb under the dowel misjudging the location of the dowel resulting in limb placement directly below the dowel or finally, a limb dropping through an empty space between the dowels. It is important to note that we evaluated the animals only in a forward in an attempt of forward movement/ progression. For example, neither placement of an animal's limb/paw onto the cardboard sidewall of the grid walk nor reaching/moving backwards for a prior rung were considered a non-grip because both failed to meet the precise non-grip criteria.

RESULTS

Tail Hang Test The E-group and N-groups' both received a score of 2 for forelimb clasping during each month of testing. On the other hand, the W-group received a score of 1 during each month of testing indicating that they did not grasp their forepaws.
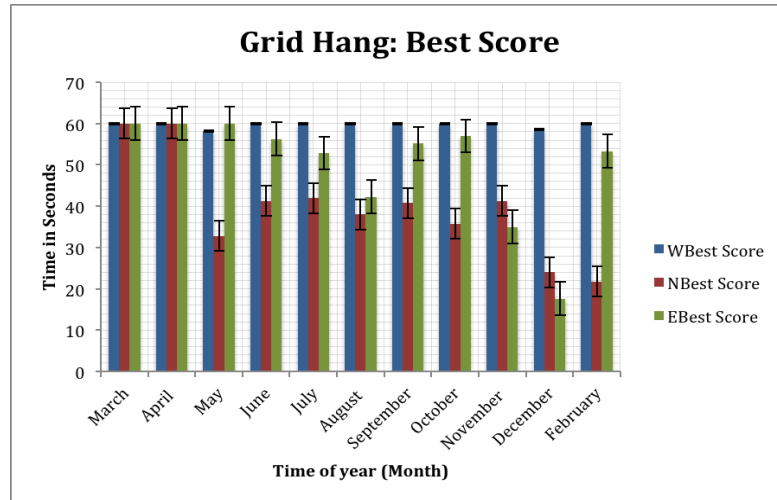
FIGURE 1. Core strength is measured through grid hang test. The results for grid hang indicate that, during the time period that the enrichment was provided (September to October), the enriched (E-group) mice hung from the grids for longer time periods than the non-enriched mice and more similar to that of wildtype (W-Group) mice. During November and December when no enrichment was provided, the E-group's grid hang score declined significantly, becoming more comparable to the nonenriched (N-Group) mice. After enrichment was reinstituted, the E-group animals' performance improved (see February) back to levels approaching the W-Group.
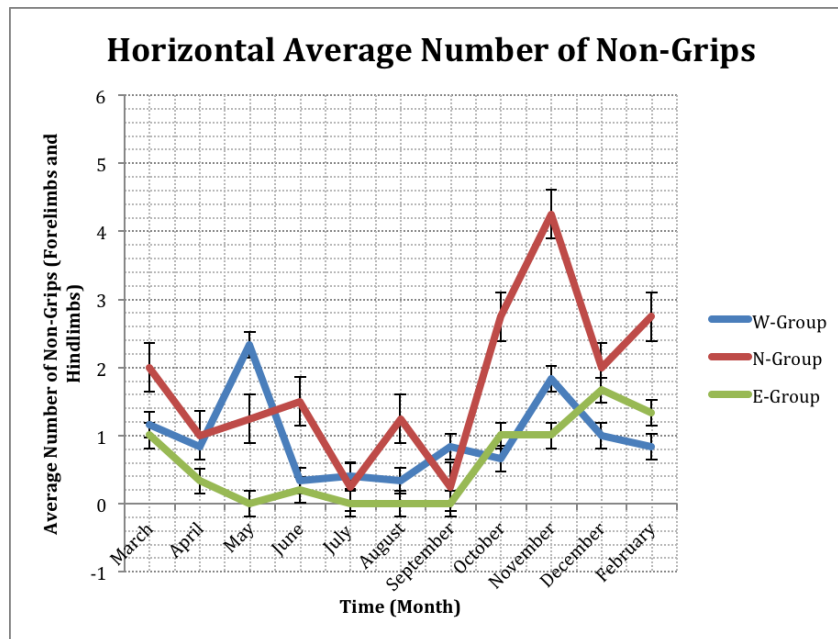

FIGURE 2. Performance on Horizontal Grid Walk, Forelimb and Hindlimb Non-Grips. While engaged in the enriching intervention, with the exception of July and September, the mice in the E-group made significantly fewer errors (less non-grips) than the N-group. Errors in the E-Group increased during the months when intervention ceased and showed improvement in February when intervention was reinstituted. While the N-Group continued to increase the number of errors as they aged, the E-group was more consistent with the W-Group.
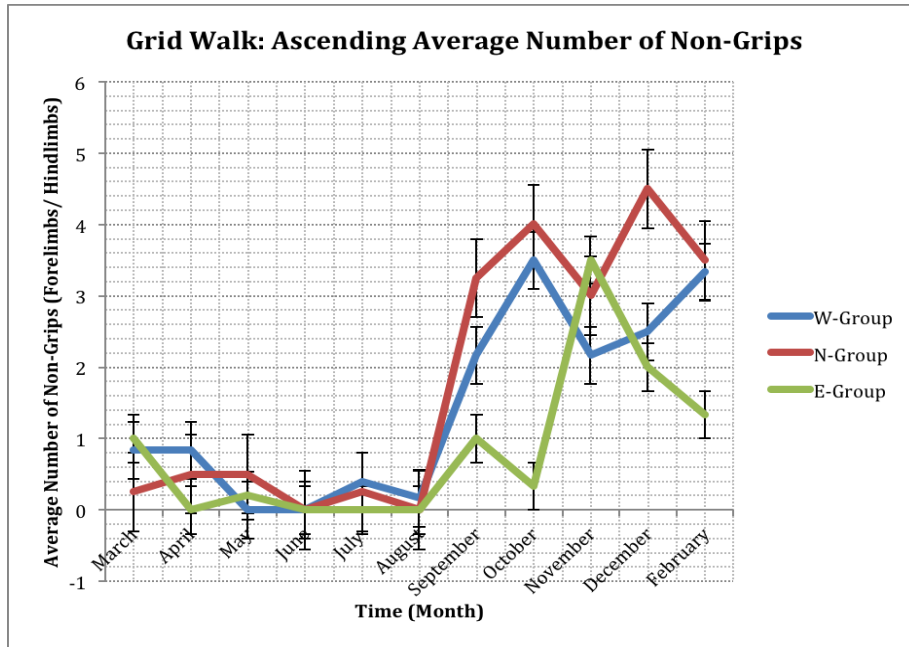
FIGURE 3. Performance on Ascending Grid Walk, Hindlimbs and Forelimbs Non-Grip Combined Score. The average number of non-grips from the months of March through February are shown. During the month of November when no enrichment was provided, the E-group mice had an unusually high average number of non-grips. By February's testing, the E-group returned to significantly fewer non-grips as compared to the N-group.
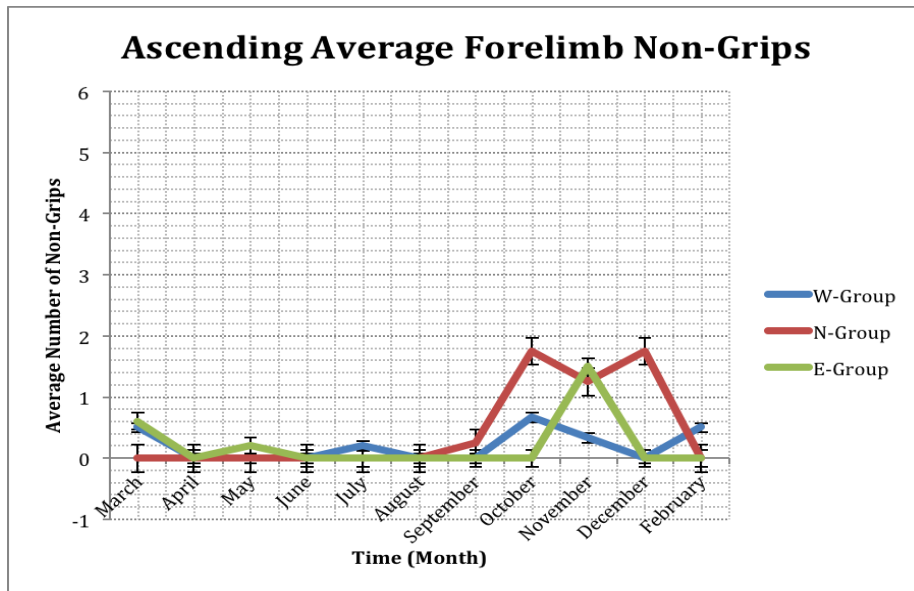

FIGURE 4. Performance on Ascending Grid Walk, Forelimb Non-Grips. Except for October through December for the N-Group and the month of November for the E-Group, most mice forelimb nongrips contributed minimally to the errors made during precise locomotion as tested by the ascending grid walk (typically receiving a score of zero nongrips).

Grid Hang Considering the size of the sample, the data was analyzed mainly through trend observation (Figure 1). To calculate all means, the best score from the three trials was used since we were interested, at this time, in what each animal's/ group's capabilities were, not necessarily how many times they performed at that level or how the fatigue factor may or may not have influenced performance. Through the first testing period (March 19-October 1) the mean times of the W-group, E-group and N-group were 59.77, 55.44 and 43.81 seconds, respectively (Figure 1). There was a significant group effect during the first testing period (F=11.29, P=0.000). The E-groups maintained their grip on the grid hang significantly longer than the N-group. However for the first month of testing, a time period before the onset of RTT symptoms, the average time spent on the grid hang for the three groups was not significantly different (Figure, 1). Another set of average scores for the best performance of the animals in each group was computed for the second testing period (October 2-December 28) when the animals did not receive any treatment ($\mu_{W-group}$=59.25, $\mu_{N-group}$=32.62, and $\mu_{E-group}$=26.33). For this period the N-group and E-group's performance were not significantly different (F=0.05, P=0.828). Finally, the last set of testing was performed in February after enrichment had been reinstituted ($\mu_{W-group}$=60, $\mu_{N-group}$=21.75, and $\mu_{E-group}$=53.33). During this testing, the performance of the E-group animals was significantly better than the N-group animals (Figure 1).

Grid Walk Trend analysis was also used to investigate the difference between the performances of the three groups in the behavioral testing of grid walk. Each animal's number of non-grips was tallied for the three subsets of tests: horizontal, ascending and descending. The results were compared between the three groups. The average rather than total number of non-grips were compared per group since the number of animals in each group differed.

A) Horizontal Grid Walk The horizontal grid walk loaded both the forelimbs and hindlimbs equally. After computing forelimb and hindlimb non-grips, the data from the 11 months was compiled (Figure 2). The observed trend shows that for the period between March and September, the N-group had in general more non-grips than the E-group animals (Figure 2). Although this trend continues in October and November, the E-group animals had significantly more non-grips than the previous

months. This is consistent with the time period where the E-group did not receive enrichment.

B) Ascending Grid Walk The ascending grid walk was used as a testing battery that loaded the hindlimbs. The average numbers of non-grips (both forelimb and hindlimb non-grips) were compared.   For the months March through August, all three groups of animals performed similarly and there does not appear to be a significant difference in their motor performance (Figure 3). Performance begins to worsen for all three groups in the months of September and October. Further, in the month of November, while enrichment was discontinued, the E group's performance was inferior to that of the N-group, noted by the increased number of failed grips than the N and W groups. But in the month of February, testing revealed that the E group had the fewest number of non-grips as compared to the N and W groups (Figure 3).

To analyze the specific performance of individual limbs, separate graphs showing the errors of the forelimbs and hindlimbs have been constructed (Figure 4 and Figure 5). The hindlimbs comparison graph shows that during the months of March through September motor abilities fluctuated.   From September to December the N-group performed worse than the period from March to September. However, except for the month of November, the E-group had fewer non-grips than the N-group suggesting that they had better limb placement (Figure 4).

C) Descending Grid Walk The descending grid walk was used as a testing battery that loaded the forelimbs. This test reveals numerous differences between the performances of the three groups.   The average number of non-grips (combined forelimb and hindlimb non-grips) shows that unlike the grid-hang test, all three groups of the animals generally had more non-grips at the beginning-testing month of April (Figure 5), which may be due to the more difficult test parameter of the downhill walk.   However, on average, over the subsequent months the E-group animals made fewer mistakes than the N-group, demonstrated by the fact that they had a lower number of non-grips compared to the N-group animals ($\mu_{E-group}$= 1.89, $\mu_{N-group}$=3.25) during the 11 testing months. To check for a change in performance after the enrichment ceased, separate means were computed for the first testing period and the second testing period ($\mu_{Period1}$=1.75, $\mu_{Period\ 2}$=2.75).   Although performing a t-test would be meaningless due
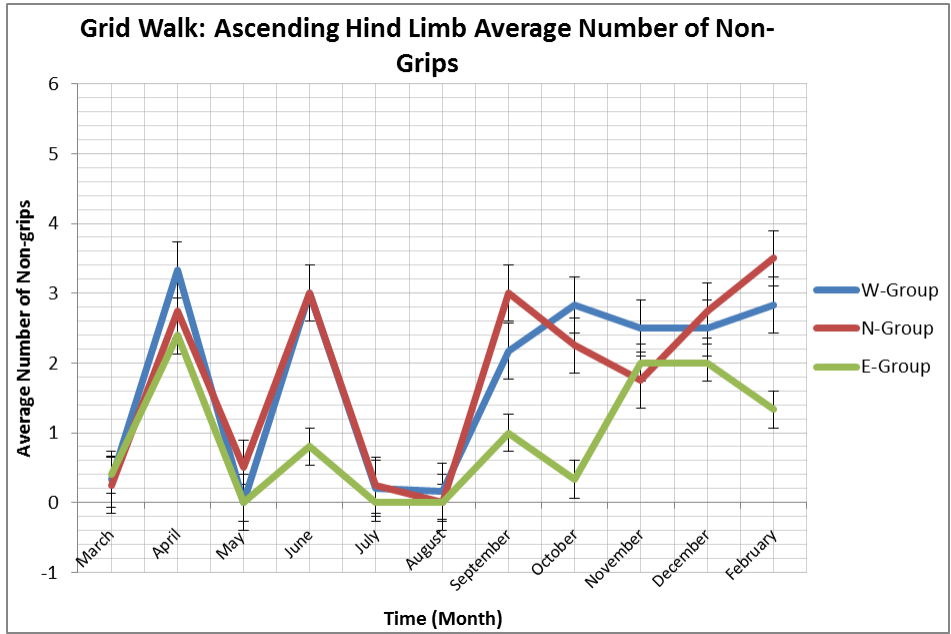
FIGURE 5. Performance on Ascending Grid Walk, Hindlimb Non-Grips. As the mice matured (September – February), the E-group experienced fewer non-grips compared to the N- or W-group. During November, when no intervention was provided, the E-Group had more difficulty with hindlimb placement (increased number of non-grips).
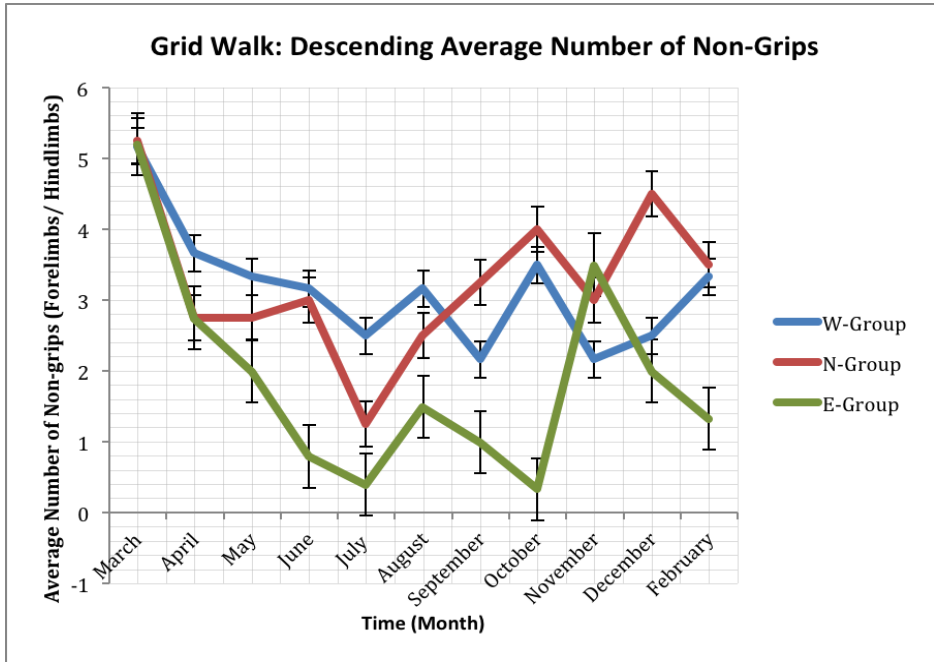

FIGURE 6. Performance on Descending Grid Walk, Forelimb and Hindlimb Non-Grips. The E-group outperformed mice in the N- and W-Group for least number of errors in grip. The one exception was during the month of November, when errors were equivalent to the N-Group. Scores returned to significantly better than all groups when enrichment was reinstituted prior to February testing.
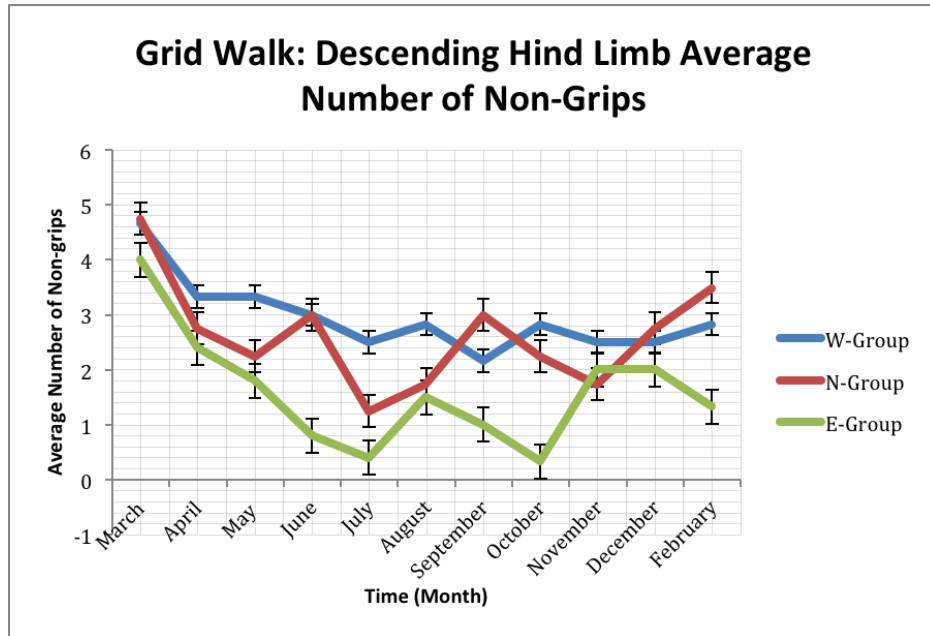
FIGURE 7. Performance on Descending Grid Walk, Hindlimb Non-Grips. The E-group mice generally exhibited significantly less hindlimb gripping errors throughout the study period. Non-grips spiked during the month of November when no enrichment was provided.
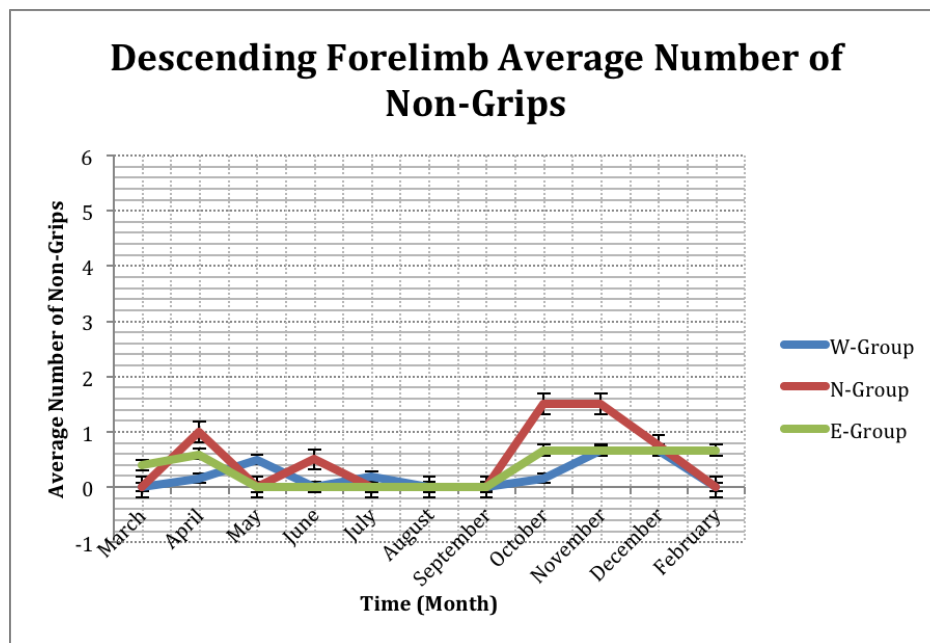
FIGURE 8. Performance on Descending Grid Walk, Forelimbs Non-Grips. All groups of mice performed similarly, generally exhibiting less than 1 forelimb gripping error throughout the testing period, with the exception of testing in October and November when the N-group made significantly more errors than the E- or W-Groups. While the E-Group made significantly more errors, on average, than the W- or N-Group in February, they continued to exhibit, on average, less than 1 non-grip.

to the small sample size, the trend suggests that E-group mouse performance worsened during the testing period when enrichment was withdrawn from the mice (Figure 5).

Further trend analysis was used to detect performance differences between the hindlimb and forelimbs when walking in a descending plane. The hindlimb analysis (Figure 7) shows that the enriched group, in general, had the least number of non-grips. The E-groups had a lower number of hindlimb non-grips compared to N-group ($\mu_{E\text{-group}}$=1.60, $\mu_{N\text{-group}}$=2.64) computed for all testing months (Figure 7). The forelimb analysis shows that, in general, for every month the E-group had lower number of non-grips compared to the N-group (Figure 8). However, the number of non-grips increased significantly for the E-group testing during the months of November and December reaching the level of the N-group during November.

DICUSSION

The results of the tail hang assessment confirmed that the Mecp2 mice possessed a defining phenotypic characteristic of forelimb grasp when held by the tail, inverted in space. This characteristic was maintained throughout the study in the two Mecp2 groups. The wild-type animals did not display these movements, again suggesting that such movements are only typical in RTT mice.

The results of the Grid Hang test suggest that enrichment helped increase the core strength of the mice. In fact, during select months, performance of the E-group and the W-group were not significantly different. This supports the hypothesis that sensorimotor enrichment can ameliorate the phenotypic behavioral symptoms of *Mecp2* null mice. It is also notable that when enrichment stops during the second testing period, the performance of the E-group animals worsened such that there was no significant difference in the months November and December in the performances between the E-group and N-group. Behavior deteriorated as early as one month after treatment ceased and recovered when the intervention was reinstituted (December 15- February 29). Since there was only one month of behavioral testing after treatment restarted (February), it is difficult to determine if a trend exists. Nonetheless, the February results suggest that a certain extent of stimulation might be needed to allow for continued progress in the maintenance of motor skills in RTT.

Kondo et al. (2008) reported an improvement in performance of the hemizygous male animal model using the standard accelerating rotarod that tests coordination and balance. Kondo et al. (2008) saw limited improvements with the rotarod while we saw generalized motor skill improvements with our more sensitive testing measures. Our results suggest that the grid hang testing apparatus may be a more sensitive testing measure for detecting subtle differences in core strength, skilled motor activity, and coordination and postural/balance reactions. Kondo et al. (2008) report housing the mice in an enriched environment but do not report monitoring the animals to ensure that they were engaging with their environment. In this study, enrichment was extensive (4hrs/day, 5times/wk.). When the animals were placed in the enriched environment, they were also monitored carefully to ensure that they were actively engaged. When lack of interaction within the environment occurred, toys were moved around to increase the activity of individual mice. Thus, the difference in the level of improvement reported by Kondo et al. (2008) and this study may be as a consequence of multiple targeted sensorimotor enriched environment and more sensitive testing measures.

Similar to the grid hang test, the results from the horizontal grid walk also indicate that enrichment may help improve RTT like symptoms. The average number of total non-grips in the horizontal grid walk test show that the enriched animals generally had fewer non-grips compared to the non-enriched *Mecp2* null mice. Moreover, the results of the performance of E-group animals on the horizontal grid walk test were significantly lower during the months of November and December when intervention ceased, reiterating the importance of continuing sensorimotor enrichment and testing for extended time periods to achieve and maintain gains in proprioceptive and limb placement abilities.

A separate analysis of the forelimb and hindlimb non-grips for the horizontal grid showed no significant difference. A possible explanation could be that the horizontal grid walk loads both forelimbs and hindlimbs equally. Hence, we only report the average number of total non-grips (forelimb+ hindlimb non-grips).

For the ascending grid walk test, it was expected that the hindlimbs would be loaded, which enhances the mouse's awareness of those limbs in space to allow for precise paw

placement. If it is true that the hindlimbs were not able to balance and handle the weight of the animal well, potentially due to poor reception of proprioception/body sense in space as appears to be the case in RTT (Kerr, 2006), we would expect to see more hindlimb non-grips compared to forelimb non-grips. The results confirmed this notion, indicating, in general, that there were more hindlimb errors. Again, our hypothesis that enrichment improves phenotypic manifestations of RTT was supported because the E-group performed better than the N-group. Although there was a fluctuation in performance in the months of March through July, the increase in the number of hindlimb errors resembles the result from the horizontal grid walk. The decline in performance in the months of November and December once more implies that when the mice are not exposed to an enriching intervention, they make more mistakes while walking up the grid walk.

For the descending grid walk test, the average number of forelimb non-grips was expected to be more than the average number of hindlimb non-grips for each animal since the design of the test loads the forelimbs more. However, the results show that on average there were more hindlimb non-grips for each group of animals. When mice were descending the grid walk, their whiskers were pointed forward, so they may have been better at detecting the position of the next grip even though they are loading their forelimbs more.

It is interesting to note that during some of the months the W-group had more non-grips than both the E-group and N-group. Although it may appear that the wild type mice are performing worse than the other groups, during the video analysis it was observed that the W-group were very quick in crossing the grid walk. In our analysis, we did not take into consideration the duration of time the animal took to cross the grid. Increased speed, however, may have compromised accuracy for the W-group of animals. Nonetheless, the comparison between the N-group and E-group is still meaningful because we did not observe considerable speed difference between these two groups.

In agreement with others (Kondo et al., 2008), the mice in this study that were allowed to actively explore enriched environments improved motorically, which suggests that enrichment is effective in multiple models of RTT. Further exploration will need to be carried out to determine if, in our model, we also see an up-regulation of BDNF as others have shown (Kondo et al., 2008).

CONCLUSIONS

The results of this study contribute to empirical findings (Kondo et al, 2008; Tabuse et al., 2010; Diener, 2002) that enrichment improves performance of multiple rodent models of neuromuscular disorders including Mecp2$^{-/y}$ mice. Moreover, the testing batteries assessed were able to identify subtle differences in performance. We argue that our assessment batteries test the mice in environments that resemble the natural places they inhabit and therefore have more functional relevance for analyzing motor recovery. It was especially enlightening to observe the animals recuperating when enrichment was reinstituted after withholding it from them in the months of October through December. Our pilot study results suggest that exposure to or active exploration within multisensory environments may be beneficial when used to treat RTT.

Nonetheless, this study has several limitations. We have primarily reported on trends since the sample size of the animals is very small. Assessing the effect of enrichment on mice should be repeated to increase the sample size and better report on the sensitivity of our testing battery to assess for motoric changes as a result of the enrichment intervention. Second, the results reported for the grid walk are preliminary data since one person has coded them. Once secondary coding is completed for all months we will reconcile scores. This would allow us to report on the reliability and validity of the novel testing protocol. Finally, the results of the testing in February suggest that the time frame for analysis of motor recovery following cessation of treatment should be extended.

References

Amir, R.E., Van der Veyver I.B., Wan, M., Tran C.Q., Francke U., Zoghbi, H.Y. (1999). Rett Syndrome is caused by mutations in X-linked

MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, *23*, 185-188.

Berrocal, Y., Pearse, D.D., Singh, A., Andrade, C.M., McBroom, J.S., Puentes, R., Eaton, M.J. (2007). Social and environmental enrichment improves sensory and motor recovery after severe contusive spinal cord injury in the rat. *Journal of Neurotrauma, 24*, 1761-1772.

Carlson, G. (2012). The use of four limb hanging wire tests to monitor muscle strength and condition over time. *TREAT-NMD.* DMD_M.2.1.005 1.0.

Chen, W.G., Chang, Q., Lin, Y., Meissner, A., West, A.E., Griffit, E.C., Jaenisch, R., Greenberg, M.E. (2003). Derepression of BDNF transcription involves Calcium-dependent phosphorylation of MeCP2. *Science, 302,* 885-889.

Cummings, B.J., Engesser-Cesar, C., Anderson, A.J. (2007). Adaptation of a ladder beam walking task to assess locomotor recovery in mice following spinal cord injury. *Behavioral Brain Research, 177*, 232-241.

Diener P.S. & Bregman B.S. (1998). Fetal spinal cord transplants support the development of target reaching and coordinated postural adjustments after neonatal cervical spinal cord injury. *Journal of Neuroscience, 18*, 763-778.

Diener, P.S. (2002), Enriched environments improve skilled movements following neonatal cervical spinal cord injury in rats, Program No. 633.20. 2002 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience. CD-ROM.

Hagberg, B., Aicardi, J., Dias, K., Ramos, O. (1983). A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Annals of Neurology, 14*, 471-479.

Kerr, A. M. (2006). Reflections on the constraints and opportunities in therapy in Rett syndrome. *The Scientific World Journal*, *6*, 992-997.

Klintsova, A.Y., Cowell, R.M., Swain, R.A., Napper, R.M.A., Goodlett, C.R., Greenough, W.T. (1998). Therapeutic effects of complex motor training on motor performance deficits induced by neonatal binge-like alcohol exposure in rats: I. Behavioral results. *Brain research, 800*, 48-61.

Kondo, M., Gray, L.J., Pelka, G.J., Christodoulou, J., Tam, P.P.L., Hannan, A.J. (2008). Environmental enrichment ameliorates a motor coordination deficit in a mouse model of Rett syndrome- Mecp2 gene dosage effects and BDNF expression. *European Journal of Neuroscience, 27*, 3342-3350.

Merkler, D., Metz, G., Raineteau, O., Dietz, V., Schwab, M.E., Fouad, K. (2001). Locomotor recovery in spinal cord-injured rats treated with an antibody neutralizing the Myelin-associated neurite growth inhibitor Nogo-A. *The Journal of Neuroscience, 10*, 3665-3673.

Moretti, P. & Zoghbi, H.Y. (2006). MeCP2 dysfunction in Rett syndrome and related disorders. *Current Opinions in Genetics and Development, 16*, 276-281.

Naidu, S. (1997). Rett syndrome. *Indian Journal of Pediatrics*, *64*, 651-659.

Pajoohesh-Ganji, A., Byrnes, K. R., Fatemi, G., & Faden, A. I. (2010). A combined scoring method to assess behavioral recovery after mouse spinal cord injury. *Neuroscience research*, *67*(2), 117-125.

Percy, A.K. (1992). The Rett syndrome: the recent advances in genetic studies in the USA. *Brain &Development*, 14, S104-5.

Tabuse, M., Yaguchi, M., Ohta, S., Kawase, T., Toda, M., (2010). A simple behavioral test for locomotor function after brain injury in mice. *Journal of Clinical Neuroscience, 17*, 1412-1416.

Tillerson & J., Miller, G. (2003). Grid performance test to measure behavioral impairment in the MPTP-treated mouse model of Parkinsonism. *Journal of Neuroscience Methods, 123*, 189-200.

Sun Y., & Wu, H. (2006). The ups and downs of BDNF in Rett syndrome. *Neuron, 49*, 321-323.

Young, J., & Zoghbi, H. (2004). X-chromosome inactivation patterns are unbalanced and affect the phenotypic outcome in a mouse model of Rett syndrome. *American Journal of Human Genetics, 74*, 511-520.

Remembering a Just World: Motivated Recall of Victim Culpability

Sahil Sharma
New York University

Research over the last 30 years has demonstrated that individuals will often blame the victim for his or her misfortune. Just World Theory (Lerner, 1980) argues that individuals do so because they are motivated to perceive their world as fair and just. Gender seems to moderate the effect of Belief in a Just World (BJW) on victim blame. Conflicting evidence suggests that this motivation affects women in different ways from men—she either blames the victim more (Janoff-Bulman, 1980) when there is a threat to the just world or less (Foley & Pigot, 2000) regardless of threat. It is less clear whether just world concerns impact recall of actual victim culpability. In this paper, we investigate whether individuals misremember information about victim responsibility for a sexual assault in order to satisfy the goal to believe that the world is just. We hypothesize that individuals whose just world motive has been experimentally heightened will be more likely to misremember details of a sexual assault in a way that confers responsibility on the victim. Results showed that memory mediates victim-blame. Men, when faced with a high threat to their belief in a just world, blamed the victim more than did women and misremembered the victimization of a female to inculcate greater blame.

Imagine you are out with your friends and you see a couple at the bar. You glance at them and then you continue to talk to your friends. Later, you receive a phone call from the police and you find out that the woman has been sexually assaulted, and that the man has not been apprehended. The officer then asks you to provide any information you can about what you observed. What would you say to the officer? If, during your attempt to recall relevant information, you thought of the way the woman was dressed or the amount of drinks it appeared she had consumed, you would not be alone. While societal norms have taught us that victims of sexual assault are *never* to blame, individuals still often wonder what the victim could have changed or should have done to prevent the victimization.

Such reactions can perhaps be explained as a rational reflection of these people's personal experiences—if in their own lives, they have avoided misfortune such as sexual assault, it might be reasonable for them to believe that the victim must have somehow contributed to their negative outcomes. However, research by Melvin Lerner suggests that this belief may not

be an entirely rational extrapolation based on personal experience. Rather, individuals may, in fact, have a *need* to believe that the world is just, and they will work to circumvent threats to that belief. According to the Belief in a Just World (BJW) theory, people want to believe that the world is fundamentally just. Maintenance of the belief in a just world implies believing that only good people receive good outcomes and only bad people receive bad outcomes (Lerner, 1980). When an unjust act occurs, such as an assault, the theory predicts that people will first try to restore justice. However, if restoring actual justice does not seem possible, they will engage in strategies such as denial or reinterpretation of the event. For example, observers have been found to restore perceived justice by attributing more negative traits and behaviors to the victims and thus blaming the victim for their injustice (Hafer, 2000).

Lerner has suggested that people differ in the extent to which they believe that the world is a just place, and that these differences can be evaluated with self-report methods. BJW has been measured by specific self-report scales, and the construct validity of these scales has

been tested against many varying situations of injustice (Hafer & Bègue, 2005).For instance, individuals may be chronically higher in the need to believe in a just world (see Hafer & Bègue, 2005, for a review). Therefore, chronically high just world believers, as indicated by these scales, are predicted to be more motivated to defend their belief in justice.

There are important moderators of belief in a just world. Evidence suggests that gender may also interact with BJW, though the research is conflicting. Kleinke and Meyer (1990) found that men with high BJW blamed rape victims more than did those with low BJW, while women did not vary by BJW levels. Similarly, Foley and Pigott (2000) found that men with high BJW awarded less to the victim in damages than did men with a low BJW. Women attributed the same level of responsibility to the plaintiff, regardless of BJW, but interestingly, those with high BJW actually awarded *more* damages to the victim. These findings may suggest that women blame female victims less because they are sympathetic to wrongs against gender in-group members. However, other research suggests that individuals are actually *more* likely to blame similar others than dissimilar others, in an effort to assuage concerns that they may suffer a similar fate to the victim's (Janoff-Bulman, 1982). In this case, women (especially those with high BJW) would actually blame rape victims *more* than will men in an effort to maintain perceived control over avoiding sexual assaults themselves.

There are also situational factors that might temporarily enhance or attenuate an individual's need to believe the world is just. One such factor that enhances the threat to the just world is when the victim's suffering continues or the perpetrator is not punished for his alleged crimes (Hafer, 2000). It is more threatening to learn that someone has victimized another and gotten away with it than to learn that they have been caught and prosecuted.

While a good deal of research has investigated the various factors affecting threat to BJW, less is understood about *how* victim blaming occurs. One such strategy that people use to blame the victim might be through the memory of the victimization. Indeed, memories are often distorted from the objective facts. People often cannot remember every small detail from an event, and thus forget numerous facts (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). Sometimes, however, such distortions of memory occur because of

motivational factors (e.g., McDonald & Hirt, 1997; Sanitioso, Kunda, & Fong, 1990). For instance, Sedikides and colleagues showed that individuals better recall self-affirming feedback compared to self-threatening feedback, apparently in order to preserve their self-esteem and identity (Sedikides, Green, & Pinter, 2004). Recent evidence suggests that motives other than self-esteem preservation, such as efforts to maintain the status quo, may also influence memory. In addition, these errors in recall appear to be specific to motivationally relevant information, and not just the result of cognitive load (e.g., Hennes & Jost, 2012).

Justice may be another motive that biases information processing, in this case directing memory towards victim blame. For example, it is possible that, after learning that a rape victim was intoxicated, an individual simply concludes that this intoxication led the victim to be assaulted. Another possibility, though, is that when people are motivated to see the world as a just place, victim blame may be facilitated by distorted information processing. For instance, the individual may be better able to recall the provocative dress of the victim, or how flirtatiously she was acting because they are so motivated to restore perceived justice.

Indeed, recent evidence suggests that people selectively recall and fabricate information to further their beliefs that people deserve their fortunes and misfortunes. For instance, Callan, Kay, Davidenko, and Ellard (2009) found that participants who were told that a lottery winner was a good person remembered the winnings as being higher than did participants who were told the winner was a bad person. They also remembered more good things about themselves after winning (versus losing) a coin flip.

As the lottery study suggests, justice motivation can even influence cognitive processing of objectively probabilistic events (such as a coin flip) for which it is clearly irrational to infer causality. The memory biases that Callan and colleagues report were not related to objective responsibility (i.e., participants were asked to recall the amount of the lottery prize rather than, for instance, the number of tickets the winner had bought). In many real-life situations, however, such as the dynamics of court cases, the attribution of blame is often a central objective. Various factors can objectively increase or decrease the likelihood of events, such as sexual assaults or car accidents, implicating responsibility as a central

consideration in analysis of these events. Therefore, the present research hopes to address how victim-blaming occurs following non-chance events.

In this study, then, we will explore the following question: how is cognitive processing of events of victimization modulated by threat to just world beliefs? We make several hypotheses in this study: First, consistent with previous research (Lerner, 1980), we predict that victim blame will be stronger when the belief in a just world has been threatened. Further, we expect this blame to be facilitated by decreased accuracy in recall of incriminating facts about the victim, and increased accuracy of exonerating information about the perpetrator. Most importantly, we expect this effect to be stronger for men, above and beyond any gender differences in chronic levels of BJW.

METHOD

Participants. 171 adult participants (80 female) were recruited from the online webservice Mechanical Turk. Participants were 78% White, and ranged in age from 18-73 ($M = 33$, $SD = 13$). Sixty-six percent of participants had at least a college education and participants were on average slightly liberal ($M = 4.25$, $SD = 2.17$), as reported on a 9-point scale ranging from extremely liberal to extremely conservative. Participants were restricted only to the United States and received $0.50 each for their participation.

Target Article: Participants were randomly assigned to one of two conditions and asked to read an article from the *Herald Times*, an online paper covering local news in Indiana. The article reports a sexual assault of a twenty-year-old female student by another male student at Indiana University. This actual article contains numerous descriptions about the victim, the perpetrator, and the crime, including clothing, GPA, academic major, descriptions about the victim's injuries, and the victim's blood alcohol level. The article was edited slightly to remove racial and other identifying information about the parties, but was otherwise largely unaltered. In the high threat to the belief in a just world ("high threat") condition, participants read that the perpetrator was still on the loose. In the low threat to the belief in a just world ("low threat") condition, the perpetrator was in police custody.

This manipulation has been found in numerous BJW studies to temporarily impact participants' motivation to believe in a just world, such that those in the "high threat" condition are more motivated to restore their belief that the world is just than are individuals in the "low threat" condition (see Hafer & Bègue, 2005, for a review).

Manipulation Check. First, participants were asked to recall if the perpetrator was caught or not.

Blame. Participants were asked the degree to which they attribute overall blame to the perpetrator or victim. Participants were asked "To what extent was the victim responsible for the assault?" and "Overall, to what extent was the offender responsible for what happened with the victim?" The participants indicated the degree to which they endorsed each statement on a 9-point scale ranging from 1 (strongly agree) to 9 (strongly disagree).

Behavior Attributions. Participants were then asked to make judgments about the behavior of both the victim and the perpetrator. They reported the extent to which participants agreed with the statements "The victim was acting promiscuously," "The perpetrator acted responsibly," and "The victim acted responsibly," on a 9-point scale ranging from 1 (strongly agree) to 9 (strongly disagree). Because responses to these items were highly correlated with each other, we created an aggregate measure of behavioral attribution ($\alpha = .76$).

Trait Attributions. Participants were also asked to indicate their degree of agreement or disagreement with three statements regarding the victim's and perpetrator's character on a 9-point scale ranging from 1 (strongly agree) to 9 (strongly disagree). Statements were, "To what extent do you feel that the victim's character was responsible for her being forced to have sex with the offender?", "The victim is a responsible person," and "The perpetrator is a violent person." These items did not demonstrate a high level of intercorrelation, so we treated each item individually in our analyses.

Recall. Participants' ability to recall details from the article was also measured. They were given seven free response or multiple-choice questions asking them to recall information attributing blame to the victim, and information

attributing blame to the perpetrator. Questions included "Do the victim's parents blame her for the assault?" and "What are the perpetrator's charges?" (see Appendix).

Belief in a Just World: Participants then completed the *Just World Scale* (Lipkus, 1991), which assesses participants' chronic need to believe that the world is a just place. Participants report their degree of agreement with seven items, such as "I feel that people get what they are entitled to have," on 6-point scales ranging from 1 (strongly disagree) to 6 (strongly agree). This scale demonstrated adequate reliability in our sample (α = .89), so scores on all seven items were averaged into a single index of Belief in a Just World.

Participants provided basic demographic information and were debriefed.

RESULTS

Manipulation Check Most participants correctly reported whether the perpetrator had been caught. Of the 21 (12%) participants who answered incorrectly, 17 were in the not caught condition and four were in the caught condition. Because the pattern of results did not differ when these participants were excluded, all data is presented here.

Victim Blame Bivariate correlations between BJW, victim blame, behavior attributions, and trait attributions are presented in Table 1.

We first assessed the interaction between gender and threat condition on victim blame. We also wanted to ensure that any gender differences observed were maintained after adjusting for chronic levels of belief in just world. Therefore, Belief in Just World was entered as a covariate. An analysis of covariance (ANCOVA) was conducted and consistent with previous research, higher scores on BJW predicted greater victim blame $F(1,166) = 12.25$, $p = .001$. There was a marginal main effect of gender, $F(1,166) = 3.29$, $p < .08$ such that men tended to blame the victim more ($M = 3.55$, $SD = 2.47$) than did women ($M = 2.61$, $SD = 2.18$). There was no main effect of condition, $F(1,166) < 1$. However, there was a significant interaction between gender and condition, $F(1,166) = 3.94$, $p < .05$. Simple effects analyses revealed that gender did not moderate victim blame when the perpetrator had been caught $t(166) < 1$. Yet when the perpetrator had not been caught, men

($M = 3.79$, $SD = 2.46$) blamed the victim to a greater extent than did women ($M = 2.12$, $SD = 1.79$), $t(166) = 2.61$, $p = .01$ (see Figure 1). Thus, particularly when the perpetrator had not been caught, men were more inclined to blame the victim for her injustice than were women.

We also evaluated whether gender and condition impacted perpetrator blame. An ANCOVA revealed that higher BJW predicted less perpetrator blame $F(1,166) = 3.74$, $p < .06$. There was also a marginal main effect of gender $F(1,166) = 2.79$, $p < .10$ such that men tended to blame the perpetrator less ($M = 7.77$, $SD = 1.89$) than did women ($M = 8.38$, $SD = 1.62$). There was no main effect of condition $F(1,166) < 1$ nor a significant interaction effect between gender and condition, $F(1,166) < 1$ (see Figure 2). This suggests that men tend to blame perpetrators less for sexual assaults than do women, but that this tendency is not affected by threat condition.

Behavior Attributions Consistent with the overall blame attribution findings, higher BJW also predicted greater victim blame through negative behavioral attributions $F(1,166) = 12.64$, $p < .001$. There was also a main effect of gender $F(1,166) = 12.39$, $p = .001$, such that men tended to blame the victim's behavior more ($M = 4.08$, $SD = 1.86$) than did women ($M = 2.86$, $SD = 1.79$). There was no main effect of condition $F(1,166) < 1$. There was a significant interaction between condition and gender $F(1,166) = 3.91$, $p = .05$. Simple effects analyses revealed that gender did not moderate victim blame when the perpetrator had been caught $t(166) = 1.19$, $p = n.s$. However, when the perpetrator had not been caught, men ($M = 4.30$, $SD = 1.84$) blamed the victim's behavior to a greater extent than did women ($M = 2.51$, $SD = 1.54$), $t(166) = 3.81$, $p < .001$ (see Figure 3). Thus, particularly when the perpetrator had not been caught, men were more inclined to believe that the victim engaged in negative behaviors that ultimately conferred blame upon her.

Trait Attributions An ANCOVA was conducted on responses to the statement, "To what extent do you feel that the victim's character was responsible for her being forced to have sex with the offender?" Findings were similar to those of victim blame and behavior attributions. Higher scores on BJW predicted greater blame on the victim's character $F(1,166) = 6.13$, $p < .02$. There was a main effect of gender $F(1,166) = 11.52$, $p = .001$ such that overall, men attribute

TABLE 1. Bivariate correlations between Belief in Just World and Victim Blame.

	Belief in Just World Scale	Victim Blame	Perpetrator Blame	Negative Behavioral Attributions	Victim Character	Victim Promiscuity	Victim Responsible Person	BAC	Reported Victim Intoxication
Belief in Just World Scale	1	.297**	-.185**	.323**	.251**	.200**	.174*		
Victim Blame	.297**	1		.734**	.689**	.536**			.175**
Perpetrator Blame	-.185*	-.331**	1	-.252**	-.363**	-.300**		.235**	.212**
Negative Behavioral Attributions	.323*	.734**	-.252**	1	.732**	.703**		-.192**	.260**
Victim Character	.251**	.689**	-.363**	.732**	1	.466**		-.223**	
Victim Promiscuity		.536**	-.300**	.703**	.466**	1			
Victim Responsible Person	.174*						1		-.197**
BAC			.235**	-.192**	-.223**			1	
Reported Victim Intoxication		.175**	.212**	.260**			-.197**		1

indicates p < .05; ** indicates p < .01; *** indicates p < .001.
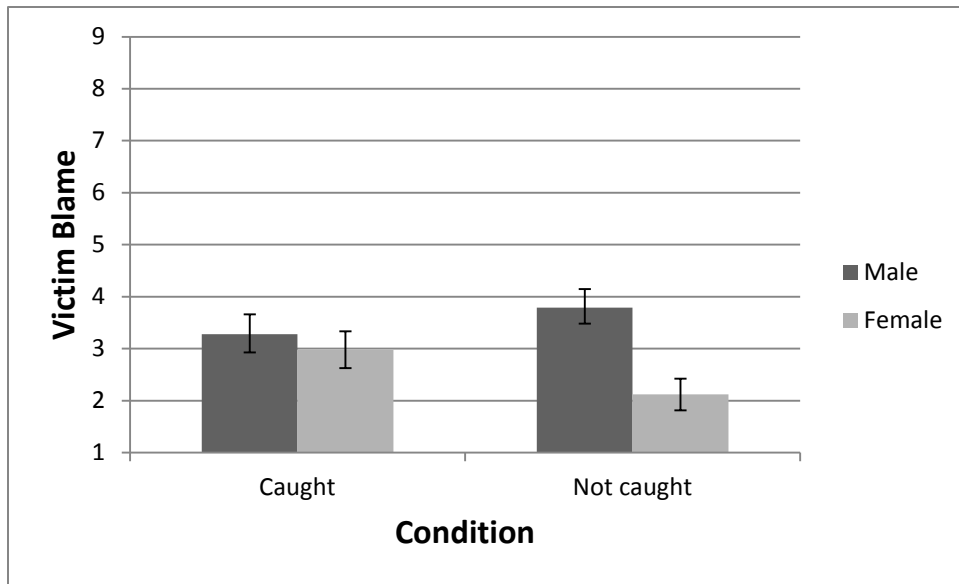
FIGURE 1. Victim Blame as a function of gender and threat condition

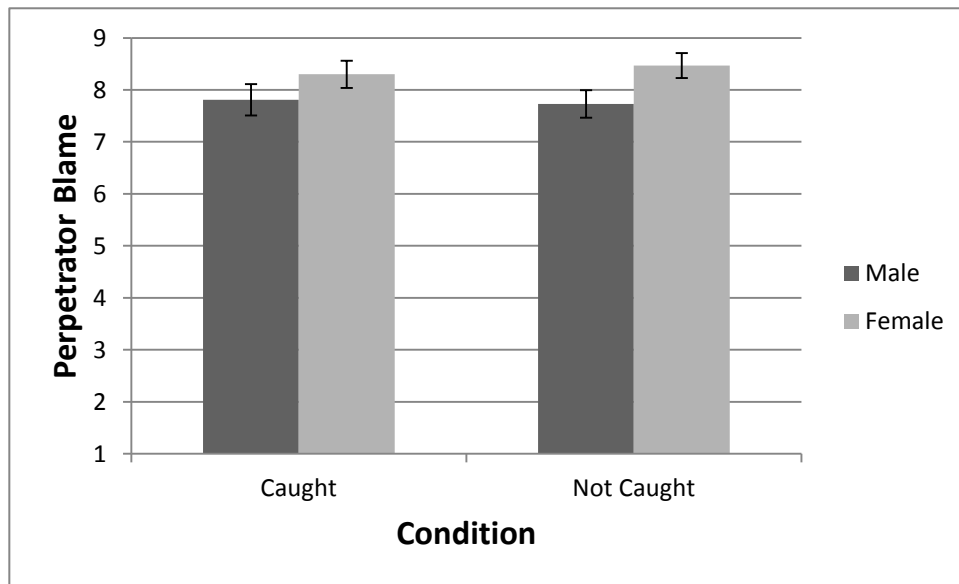FIGURE 2. Perpetrator Blame as a function of gender and threat condition
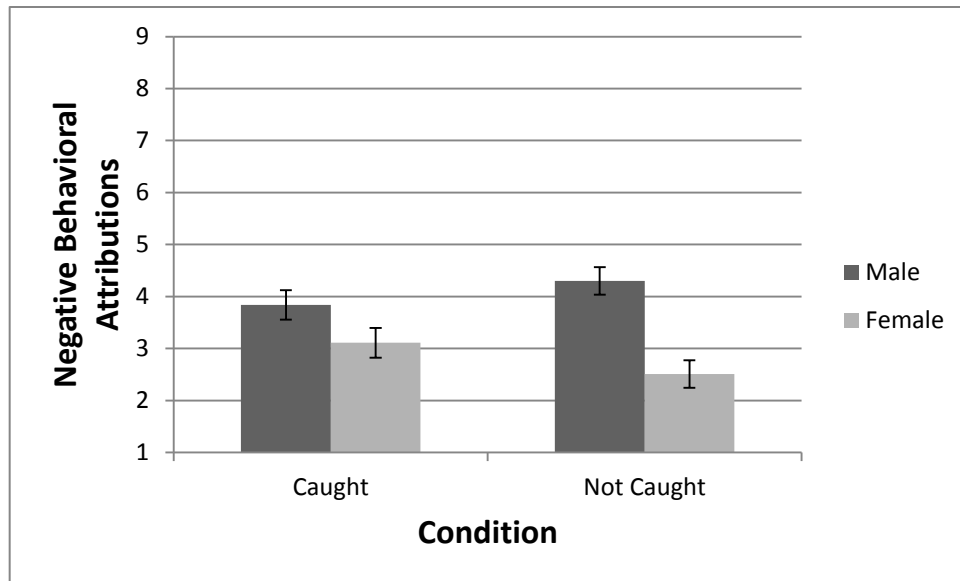
FIGURE 3. Negative Behavior Attributions as a function of gender and threat condition
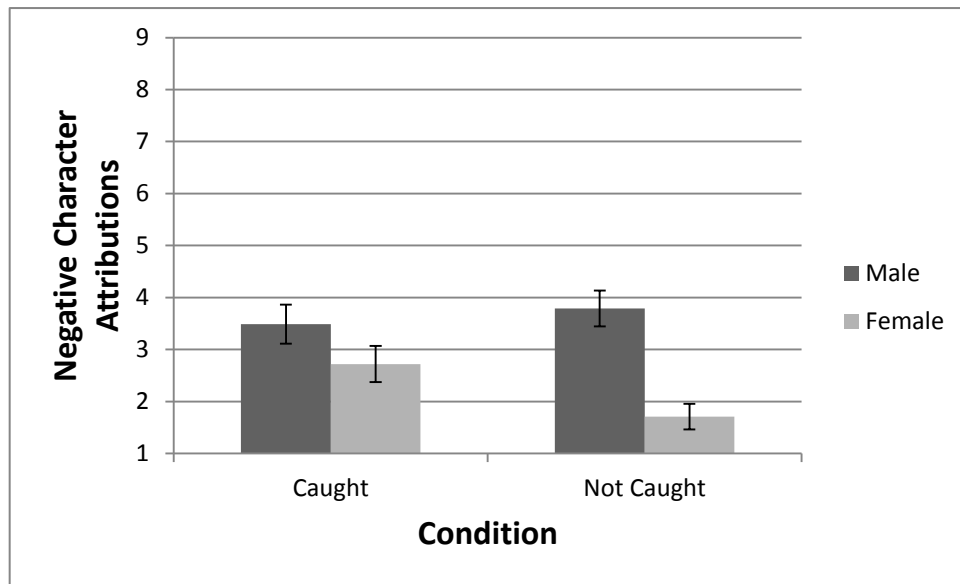

FIGURE 4. Negative Trait Attributions as a function of gender and threat condition
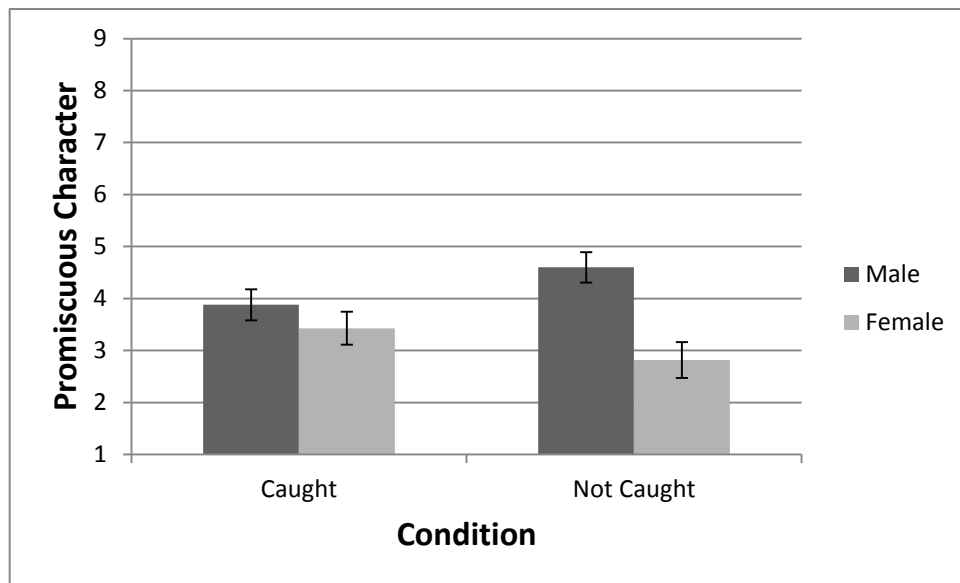
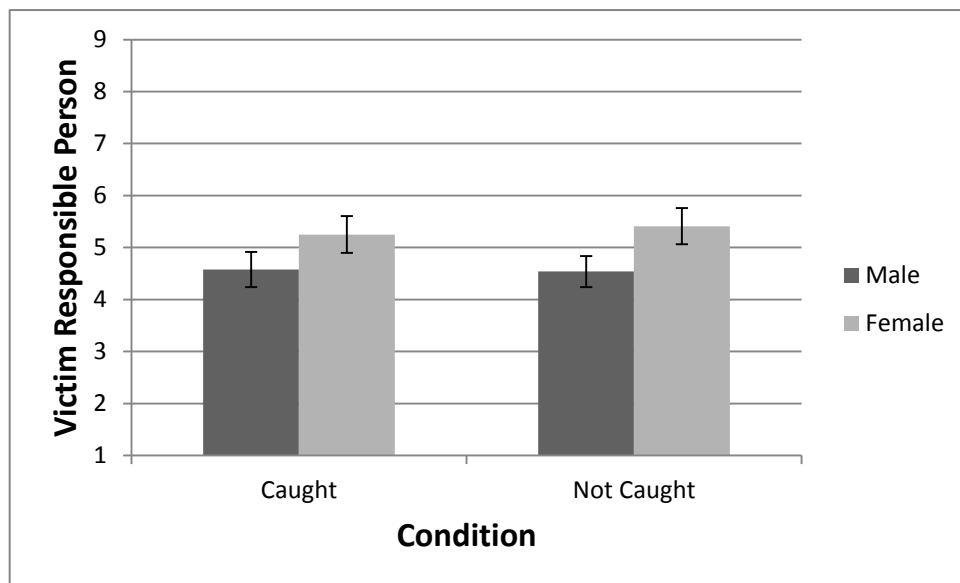FIGURE 5. Promiscuous Character as a function of gender and threat condition

FIGURE 6. Victim Responsible Person as a function of gender and threat condition

TABLE 2. Distribution of responses to the question, "Was the victim conscious during the incident?"

	Conscious	**Not Conscious**
Male Not Caught	15%	**85%**
Female Not Caught	0%	**100%**
Male Caught	19%	**81%**
Female Caught	13%	**87%**

$\chi^2(3) = 10.72$, $p < .02$ (Percentages are within gender by condition. Correct answers are in bold.)

TABLE 3. Distribution of responses to the question, "What are the perpetrator's charges?"

	Less Strict	**Correct Charge**	Most Strict
Male Not Caught	29%	**52%**	19%
Female Not Caught	18%	**82%**	0%
Male Caught	28%	**67%**	5%
Female Caught	30%	**67%**	2%

$\chi^2(9) = 22.02$, $p < .009$ (Percentages are within gender by condition. Correct answers are in bold.)

more negative character traits to the victim ($M = 3.65$, $SD = 2.41$) than did women ($M = 2.29$, $SD = 2.07$). There was no main effect of condition $F(1,166) < 1$, but there was a marginal interaction between condition and gender $F(1,166) = 3.76$, $p < .06$. Simple effects analyses showed that gender did not moderate victim blame when the perpetrator had been caught $t(166) = 1.13$, $p = n.s.$. However, when the perpetrator had not been caught, men ($M = 3.79$, $SD = 2.38$) blamed the victim's character to a greater extent than did women ($M = 1.71$, $SD = 1.43$) $t(166) = 3.69$, $p < .001$. (see Figure 4). Thus, men in the not caught condition made significantly more derogatory trait attributions upon the victim than did women.

ANCOVA analyses were also conducted to assess how participants responded to the statement, "the victim is a promiscuous person".

Higher scores on BJW predicted greater blame on the victim's character $F(1,166) = 3.66$, $p < .06$. There was a main effect of gender $F(1,166) = 8.82$, $p < .01$ such that overall, men tended consider the victim to be more promiscuous ($M = 4.26$, $SD = 1.97$) than did women ($M = 3.18$, $SD = 2.11$). There was an interaction between condition and gender $F(1,166) = 4.66$, $p < .05$. Simple effects analyses revealed that gender did not impact victim blame when the perpetrator had been caught $t(166) < 1$. However, there was a significant difference for gender when the perpetrator was not caught $t(166) = 3.54$, $p = .001$. When the perpetrator was not caught, men ($M = 4.60$, $SD = 2.02$) attributed promiscuous traits to the victim more than did women ($M = 2.82$, $SD = 2.01$) (see Figure 5). Thus, particularly when the perpetrator had not

been caught, men were more inclined to believe that the victim was a promiscuous person.

We also evaluated how participants responded to the statement, "The victim is a responsible person". Higher scores on BJW predicted stronger beliefs that the victim was not a responsible person $F(1,166) = 9.04$, $p = .003$. There was a main effect of gender $F(1,166) = 8.21$, $p = .005$ such that men ($M = 4.56$, $SD = 2.13$) thought the victim was less of a responsible person than did women ($M = 5.25$, $SD = 2.24$). There was no main effect of condition $F(1,166) < 1$ nor a significant interaction effect between gender and condition, $F(1,166) < 1$. Simple effect analysis showed that gender marginally impacted victim blame when the perpetrator had been caught $t(166) = 1.78$, $p < .08$, such that men thought the victim was not a responsible person ($M = 4.58$, $SD = 2.22$) to a greater extent than did women ($M = 5.13$, $SD = 2.40$). When the perpetrator was not caught, men thought the victim was not a responsible person ($M = 4.54$, $SD = 2.07$) to a greater extent than did women ($M = 5.41$, $SD = 2.03$), $t(166) = 2.33$ $p = .021$ (see Figure 6). Therefore, men were more likely to dismiss the victim as being irresponsible when the perpetrator had not been caught.

Recall Consistent with our predictions, for the recall of victim's state of consciousness, a chi square test showed a significant interaction of gender and condition $\chi^2(3) = 10.72$, $p < .02$. Follow up analyses revealed that when the perpetrator had been caught, there was no effect of gender, $p = n.s.$ However, within participants who learned that the perpetrator had not been caught, 85% of men remembered the

victim was unconscious compared to 100% of women, $p < .05$. (See Table 2).

Consistent with our predictions, a chi square test demonstrated a significant interaction of gender and condition on memory of the perpetrator's charges, $\chi^2(9) = 22.02$, $p = .009$. Follow up analyses revealed that when the perpetrator had been caught, gender did not moderate memory of the charges, $p = n.s.$, but when the perpetrator had not been caught, males (52%) were significantly less likely to correctly remember the perpetrator's charges then were women (82%). However, of those who answered incorrectly, men were significantly more likely than women to remember a *greater* charge (19% vs. 0%), $p < .05$ (See Table 3).

There was not a significant omnibus effect of gender and condition on how likely participants were to misremember the victim's clothing. $\chi^2(9) = 10.24$, $p < .12$. However, in line with our predictions, follow-up analyses revealed that within participants who learned that the perpetrator had not been caught, men (38%) were significantly less likely than women (62%) to correctly remember the victim's outfit, $p < .05$. There was no effect of gender on memory for participants who learned the perpetrator had been caught, $p = n.s.$ However, for those who answered incorrectly in the perpetrator not caught condition, men were significantly *more* likely (35%) than women to remember that she was dressed *more* conservatively (12%), $p < .05$.

There was also no omnibus interaction of gender and condition on the victim's injuries, $\chi^2(6) = 6.12$, $p = n.s.$ However, follow-up analyses

TABLE 4. Distribution of responses to the question, "What was the victim wearing?"

	1	**2**	3
Male Not Caught	35%	**38%**	27%
Female Not Caught	12%	**62%**	27%
Male Caught	16%	**49%**	35%
Female Caught	15%	**50%**	35%

$\chi^2$2(9) = 10.24, p < .12 (Percentages are within gender by condition. Correct answers are in bold. 1=conservative clothing; **2=accurate clothing**; 3=least conservative clothing)

TABLE 5. Distribution of responses to the statement, "Indicate all of the injuries the victim sustained in the assault".

	1	**2**	3
Male Not Caught	13%	**83%**	4%
Female Not Caught	9%	**88%**	3%
Male Caught	21%	**70%**	9%
Female Caught	11%	**87%**	2%

χ^2(6) = 6.12, *p = n.s* (Percentages are within gender by condition. Correct answers are in bold. 1=minor injuries, **2=accurate charge injuries**, 3=severe injuries.)

TABLE 6. Distribution of responses to the question, "Do the victim's parents blame her for the assault?"

	Blame	No Blame
Male Not Caught	**54%**	46%
Female Not Caught	**74%**	27%
Male Caught	**42%**	58%
Female Caught	**63%**	37%

χ^2(3) = 8.81, *p* < .04 (Note: Percentages are within gender by condition. Correct answers are in bold.)
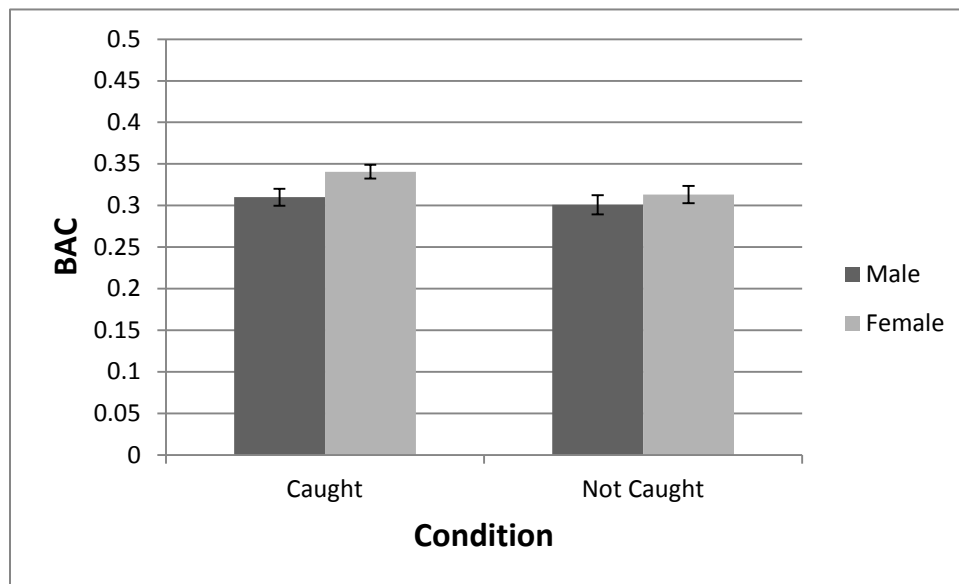
FIGURE 7. Memory of BAC as a function of gender and threat condition

revealed that within participants who learned that the perpetrator had been caught, men were less likely (70%) than women (87%) to correctly remember the victim's injuries, $p < .05$. There was no effect of gender on memory for participants who learned the perpetrator had *not* been caught, $p = n.s.$

A chi square test showed a significant interaction effect of gender and condition on memory of parental blame, $\chi^2(3) = 8.81$, $p < .04$. Inconsistent with our predictions, when the perpetrator had not been caught, gender did not moderate memory of parental blame, $p = n.s.$ However, when the perpetrator had been caught, men were less likely to correctly remember that the victim's parents blamed her (42%) than were women (63%), $p < .05$ (see Table 4).

Surprisingly, an ANCOVA revealed that BJW did not have a significant impact on memory of the victim's Blood Alcohol Content, $F(1,166) = 1.50$, $p = n.s.$ There was a marginal effect of gender $F(1,166) = 2.81$, $p < .10$, but men tended to remember the victim's BAC as *lower* ($M = .31$, $SD = .07$) than did women ($M = .33$, $SD = .06$). There was also a marginal effect for condition $F(1,166) = 3.40$, $p < .07$, but in contrast to our hypothesis, participants who were in the perpetrator not caught condition ($M =$

.31, $SD = .07$) misremembered the BAC to be *lower* than participants in the perpetrator caught condition ($M = .33$, $SD = .06$). There was no interaction between condition and gender, $F(1,166) < 1$. Simple effects analyses revealed that when the perpetrator was not caught, there was no gender difference in recalled BAC, $t(166) < 1$. However, when the perpetrator was caught, men recalled the BAC to be lower ($M = .31$, $SD = .07$) than did women ($M = .34$, $SD = .06$), $t(166) = 1.93$, $p < .07$. This suggests that men tended to remember the victim as being *less* intoxicated than did women, especially when the perpetrator had been caught.

An ANCOVA showed that higher scores on BJW predicted greater reported intoxication $F(1,166) = 6.97$, $p < .01$. There was also a main effect of gender, $F(1,166) = 4.27$, $p = .04$, but men remembered the victim to be *less* intoxicated ($M = 7.66$, $SD = 2.04$) than did women ($M = 8.42$, $SD = 1.52$). There was no main effect of condition $F(1,166) < 1$ nor an interaction effect of gender and condition, $F(1,166) = 2.24$, $p = n.s.$ Simple effects analyses revealed that when the perpetrator was not caught, men thought the victim was less intoxicated ($M = 7.48$, $SD = 2.14$) than did women ($M = 8.68$, $SD = .84$), $t(166) = 2.47$, $p < .02$, but there was no gender difference when

the perpetrator had been caught, $t(166) < 1$. Therefore, in the not caught condition, men tended to remember that the victim was *less* intoxicated than did women.

DISCUSSION

We hypothesized that individual's situational increases in the motivation to believe that the world is just would lead to increased victim blame, and that this blame would be facilitated by greater recall of incriminating facts about the victim. We also expected these effects to be more apparent among men. Indeed, this study demonstrated that men were more likely to blame the victim of a sexual assault, and less likely to blame the perpetrator, especially when there was a high threat to their BJW. Our findings suggest that the motivation to defend the just world led men in particular to perceive the victim in a more negative light, derogating both her character traits and her behaviors. Men believed the victim to be more promiscuous and her behaviors to be more responsible for the assault when the perpetrator was not caught. Extending findings by Foley and Pigot (2000), men were more affected by the threat to BJW, and higher threat led to greater belief in victim culpability, while there often was no significant difference in blame attribution among women in under both conditions. This is consistent with previous research that suggests that people are more sympathetic to the plights of similar others than to dissimilar others.

Along with reporting overall higher victim culpability, men cognitively distorted information from the news story to ultimately confer blame on the victim in the high threat condition. Men were more likely than women to misremember that the victim was conscious, perhaps to facilitate the perception that she consented to the sexual encounter. However, some recall findings were inconsistent with our predictions. For instance, men misremembered the perpetrator as having more severe charges than he actually received. However, misremembering charges may have been a way for men to convey that the perpetrator was unfairly charged. We also found that when men misremembered the victim's outfit, they actually recalled her dressing more conservatively rather than more provocatively than she was actually dressed. This finding might indicate that when men face a threat to the just world, they are more inclined than women to blame the victim

via her character or behavior, rather than clothing. Overall, while men blamed the victim more when there was a greater threat to the just world, our findings regarding memory biases were mixed. Specifically, while in some situations memory errors were consistent with more victim blame, in other situations they were consistent with less. It may be that learning that an in-group member (another male) has assaulted someone is also threatening, and this threat leads to poorer information processing. This possibility should be investigated in future research.

Although our experiment provides some support that justice motives affect recall, further studies should investigate whether biased information processing occurs during encoding or during the retrieval of memories. When exactly do people distort information and how can this distortion be remedied? In addition, it is unclear from our research whether the desire to blame the victim leads people to misremember information, or whether biased memory processing leads individuals to conclude that the victim was to blame. Perhaps participants could be asked to recall questions after a longer period of time after reading the news story. The questions asked in this study may not have been challenging enough for participants or may not have been worded to precisely measure blame, and instead may have gauged other sentiments aside from derogation.

Further studies may also investigate other paradigms of injustice, aside from assault, to test whether the gender effects seen here persist. For instance, the nature of sexual assault may have evoked more prominent differences in men and women's reactions toward the victim, especially since the victim was female. A car accident, perhaps, could serve as an additional paradigm to evaluate whether men and women differ in blame attribution. It is unclear whether women blame sexual assault victims less because they are similar to them and they are empathizing, or whether women simply blame any victim less than men do. Further, even though many researchers study BJW via vignettes in which participants are not directly involved, effects might be shown more acutely when participants themselves are eyewitnesses to the just-world-threatening event.

Despite these limitations of this study, our findings have suggestive implications for the courtroom. How reliable are eyewitness testimonies? How can eyewitness testimonies

be made more reliable and accurate? Our research suggests that eyewitnesses may misremember information from a crime given what they know about the victim's suffering. Continued suffering, such as when the perpetrator is not caught, creates a greater threat to the witness's just world. Thus, their cognitive processing of the event may be affected by their motivation to re-instate a just environment. Eyewitnesses may exaggerate details conferring blame on the victim rather than providing objective facts about the situation. Furthermore, our research suggests that during *voir dire,* attorneys may consider the effects of gender on victim blame before selecting potential jurors. These findings can help lawyers restore justice to victims who must defend themselves when a juror's sense of justice has been threatened.

This study may also be helpful during Truth Commissions, which are held to reveal crimes committed by government officials or even non-state actors. These commissions are invoked to issue retribution for human rights violations committed by the government against citizens of that nation (Hayner, 1994). The unrest is analogous to threats to the just world, insofar as they inflict mass injustice on the citizens. But how are these officials chosen for prosecution, and what evidence provided by witnesses is deemed admissible? By better understanding the mechanisms through which people blame victims, we may become better equipped to remedy, rather than exacerbate, injustice to individuals who suffer at the hands of others.

References

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 570-585.

Callan, M. J., Kay, A. C., Davidenko, N., & Ellard, J. H. (2009). The effects of justice motivation on memory for self- and other-relevant events. *Journal of Experimental Social Psychology*, *45,* 614-623.

Correia, I., & Vala, J. (2003). When will a victim be secondarily victimized? The effect of observer's belief in a just world, victim's innocence and persistence of suffering. *Social Justice Research, 16,* 379-400.

Foley, L.A. & Pigott, M. A, (2000), Belief in a just world and jury decisions in a civil rape trial. *Journal of Applied Social Psychology*, *30*, 935-951.

Hafer, C. L. (2000). Do innocent victims threaten the belief in a just world?: Evidence from a modified Stroop task. *Journal of Personality and Social Psychology, 79,* 165-173.

Hafer, C. L., & Bègue, L. (2005). Experimental research on Just-World Theory: Problems, developments, and future challenges. *Psychological Bulletin, 131,* 128-167.

Hennes, E. P., & Jost, J. T. (2012). Motivated information processing in the service of the system: The case of anthropogenic climate change. Unpublished manuscript.

Janoff-Bulman, R. (1982). Esteem and control bases of blame: "Adaptive" strategies for victims versus observers. *Journal of Personality*, *50,* 180-192.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion.* Perspectives in Social Psychology. New York: Plenum Press.

McDonald, H. E., & Hirt, E. R. (1997).When expectancy meets desire: Motivational effects in reconstructive memory. *Journal of Personality and Social Psychology, 72,* 5-23.

Rubin, Z. & Peplau L. A. (1975) Who believes in a just world? *Journal of Social Issues, 31,* 65-89.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memory. *Journal of Personality and Social Psychology, 59,* 229-241.

Sedikides, C., Green, J. D., & Pinter, B. (2004). Self-protective memory. In D. R. Beike, J. M. Lampinen, & D. A. Behrend (Eds.), *The self and memory* (pp. 161–179). New York, PA: Psychology Press.

Appendix

Recall Questionnaire

1. Has the perpetrator been caught?
    a. Yes
    b. No

2. What was the victim's blood alcohol content?
    0 - 0.05 - 0.09 - 0.14 - 0.18 - 0.23 - 0.27 - 0.32 - 0.36 - 0.41 - 0.45 - 0.5

3. How intoxicated was the victim at the time of the assault? (0, not at all intoxicated; 9, extremely intoxicated)
4. What was the victim wearing?
    a. Blue sweater, black knee-length skirt, and flats
    b. Black tank top, blue jeans, and flats
    c. Black shirt, blue knee-length skirt, and high heels.
5. Indicate all of the injuries the victim sustained in the assault.
    a. Dislocated shoulder, cuts on face, severe vaginal injuries
    b. Bite marks, swollen buttocks, severely bruised knees, cracked rib
    c. Mild bruising, shoulder pain, mild vaginal injuries

6. What are the perpetrator's charges?
    a. Three counts of sexual deviate conduct, three counts of rape, one attempted murder charge
    b. One count of rape, one count of deviate sexual conduct, one count of criminal confinement
    c. DUI, one count of sexual deviate conduct
    d. Three counts of rape, three counts of deviate sexual conduct, one count of criminal confinement

7. Do the victim's parents blame her for the assault?
    a. Yes
    b. No
8. Was the victim conscious during the incident?
    a. Yes
    b. No

An Exploration of the Negative Effects of Repetition and Testing on Memory

S. Adam Smith
The University of North Carolina at Chapel Hill

A fundamental principle within human memory research is the idea that repetition (i.e. multiple presentations of a stimulus) and testing (i.e. preliminary recall tasks) both improve recall performance. However, recent evidence suggests that in certain conditions repetition and testing can actually *decrease* item recall (Peterson, 2011). This study sought to determine whether these negative effects of repetition and testing would be more appropriately accounted for in the context of an encoding explanation or a retrieval explanation – in other words, whether the cause for decreased performance was related to how efficiently items were encoded or how effectively relevant self-cues were used during the recall task. Two experiments were designed to test these explanations by using lists of rhyming cue-target word pairs (e.g. "Beg – Leg") as stimuli. The target words of these pairs were organized pseudo-randomly in some phases and categorically in others. The ordering of these phases was intended to direct what relational information would be most salient – with initial pseudo-randomized ordering, within-pair (rhyming) similarities should be more apparent, and with initial categorical ordering, between-pair (categorical) similarities should be more easily noticed. Results of the experiments support an encoding account for the negative repetition effect, but a retrieval explanation for the negative testing effect.

A well-known proverb concerning the achievement of excellence is "practice makes perfect." Though few would truly assert that literal perfection can be attained through repetition, it is commonly accepted that repetition boosts one's performance on a wide variety of tasks. Johnstone, Ashbaugh, and Warfield (2002) were even able to demonstrate improvement in the development of writing skills – a complex cognitive task – merely through repeated practice. This use of repeated practice has been particularly prevalent in pedagogical settings, wherein the drill-and-practice technique (going over information until it is mastered) is a commonplace procedure in the teaching of academic skills – particularly those of mathematics and grammar. In a study by Brophy (1986), this strategy was especially effective among students with lower academic achievement, which lends some merit to the frequent use of the repetition-based technique. Furthermore, Brophy also observed that the usefulness of this approach is not confined exclusively to the instruction of basic skills, suggesting instead that a structured environment can be applied to "…any body of knowledge or set of skills that has been sufficiently well organized" (p. 1076). Though the idea behind the benefits of repetition is an intuitive one, this does not mean that it is void of scientific grounding. Simply put, the repetition effect merely posits that increased frequency and exposure to stimuli increases the later recall of said stimuli.

Indeed, the assertion that repetition impacts recall appears so self-evident that it hardly seems worthy of mention. However, it is worth noting because this principle alone cannot account for differences in individuals' levels of recall in practical settings. For instance, two students may spend the same amount of time studying in preparation for a test, but this in no way ensures identical performance on the exam. One must also take into account the various mnemonic strategies that can be implemented to increase the effectiveness of repetition as a

learning tool.   For instance, it has been observed that information is better retained if practice is temporally distributed (i.e. "spaced rehearsal") than if it is presented with higher frequency during a shorter time interval (i.e. "massed rehearsal).   In a study by Dempster (1987), researchers found that the benefits of this spacing effect apply to subjects who are actively attempting to learn new and unfamiliar vocabulary.   This finding is particularly compelling because it displays the utility of the spacing effect not only for memorizing familiar stimuli, but also for learning new material, thereby bolstering the assertion of its usefulness as a pedagogical tool.

In more recent years, researchers have been looking into a somewhat less explored phenomenon known as the "testing effect." Although tests are typically used to gauge the retention of knowledge, there is strong evidence to suggest that testing itself actually alters memory traces and affects later recall.   More precisely, the testing effect refers to the observation that testing augments a participant's retention of information more effectively than simply restudying the material.   In a review of the testing effect, Roediger and Karpicke (2006) further subcategorize this phenomenon into two forms: a mediated testing effect and a direct testing effect.   Mediated testing effects are those which indicate that, "it is not the act of taking the test itself that influences learning, but rather the fact that testing promotes learning via some other process or processes." (Roediger & Karpicke, 2006, p. 182).   In contrast, a direct effect of testing would attribute the increased retention of information with the act of taking a test itself rather than some alternative mediating process.

The general result of experiments which explore this effect is that a group given an initial pretest outperforms a control group (with no initial pretesting) on a final measurement of information recalled.   This effect persists even in instances where there is no feedback provided after initial testing, decreasing the likelihood that this recorded improvement is actually due to some mediating factor caused by such feedback. For the purposes of the current study, reference to the testing effect will specifically indicate this direct form of the effect.

There is another effect that is similar in nature (and results) to the testing effect known as the generation effect.   The generation effect refers to the observation that generating information from past knowledge typically results in greater memory retention than simply reading the same material.   In an experiment designed to test this effect, subjects are prompted to generate word of interest when given a meaningful cue. For instance, if the cue was based on antonyms, participants might be presented with the stimulus "hot-c___", and be expected to generate the word "cold" as a response to this cue (Mulligan & Lozito, 2004). Typically, a generation condition yields better performance on subsequent memory tests as compared to a control condition in which participants are merely instructed to read the word pairs. Although this is comparable to the testing effect, it is important to distinguish these two phenomena.   Chiefly, it should be noted that the testing effect occurs as a result of a participant accessing his or her episodic memory in order to bring to mind an item which was depicted during prior study.   In contrast, the generation effect is prompted when a target item is being generated from semantic memory in relation to a given cue (in other words, not retrieved from a specific study phase, but rather generated based on general knowledge).

The negative repetition effect and negative testing effect The positive effects of utilizing both repetition and testing to improve memory are quite well documented. However, it is important to determine if these effects are always positive. For instance, might it be possible that repetition of verbal stimuli could produce a *negative* effect on memory in certain conditions? At first glance this proposition seems not only counterintuitive but also unlikely, as it appears incongruent with the majority of memory research that has been conducted thus far. Nevertheless, despite the apparent consensus concerning positive effects of repetition on memory storage and retrieval, evidence that may call some aspects of the well-established repetition effect into question has recently surfaced.

In Daniel Peterson's (2011) recent dissertation, a number of experiments were conducted in order to determine how the "item-specific" versus "relational" account could explain certain elements of the testing effect. In short, this account holds that informational qualities of a set of stimuli are processed in two basic ways: by attending to features which are unique to a particular stimulus (i.e. item-specific processing), or by attending to features which are commonly expressed by a set of stimuli (i.e. relational processing) (Hunt & McDaniel, 1993). In interpreting the data from these experiments,

Peterson discovered an unpredicted result. Namely, participants in one condition were presented with a list of words twice, and yet they recalled 13% fewer target words than subjects in another condition who were given the same list only once. This finding is particularly surprising considering that the group given the word list twice experienced each stimulus in a spaced manner – a presentation that should have maximized the advantages of repetition. Peterson referred to this result as a "negative repetition effect".

Considering this finding, it seems reasonable as well to question whether the testing effect is always positive. Indeed, in this same dissertation, Peterson (2011) offers evidence that contradicts this invariably positive outlook on the testing effect. However, unlike the negative repetition effect, evidence for the negative testing effect was a predicted outcome of the study. The reason for this conjecture is Peterson's hypothesis that the item-specific versus relational account, which has been used to explain the similar generation effect, can explain the testing effect. Peterson surmises that because the generation effect can be shown to negatively affect memory under certain conditions, if a similar experiment were to be conducted with the testing effect replacing the generation effect, the observation of a negative testing effect would strengthen the idea that both of these phenomena can be explained via the item-specific versus relational account. In an effort to test this prediction, Peterson modeled Experiment 4 of his study after a design implemented by Burns (1990). In the study, the stimuli consisted of 36 rhyming cue-target word pairs (e.g. "Beg – Leg"). Notably, the target words of these pairs fell into one of six distinct categories; using the above example, "Parts of the Human Body" would be the target word category.

This experiment consisted of two conditions and three phases. The first condition was a "restudy" condition, wherein the participant was instructed to read through a list of word pairs two times. The second condition was a "retrieval" condition, in which the participant read through the word pair list once for the first phase, but was later asked to recall the target word of each individual pair when presented with the cue word. By comparing the performance of the retrieval condition with that of the restudy condition, Peterson would be able to determine what effect testing had on the recall of target words, whether positive, negative, or null.

In phase one of the study, subjects in both conditions were presented with the cue-target word pairs in pseudo-random order, so that no two pairs with target words from the same category appeared sequentially. In phase two, the word pairs were organized by category so that pairs with target words from the same category occurred in sequence. For the restudy condition, both words in the pair were simply presented an additional time (albeit in categorical order), and for the retrieval condition the cue word was presented in order to prompt recall of the corresponding target word. Phase three was a free recall test of the target words from the earlier pairs and was the same for both conditions. As predicted, the implementation of randomized word pairs before phase two caused the retrieval condition to perform more poorly than the restudy condition, indicating a negative testing effect.

Peterson believed that the observed negative testing effect could be explained when put in the context of the item-specific versus relational account. He asserted that when stimuli are deemed unusual (in this instance, an incomplete word pair that requires retrieval of a target word), more attention is allocated to each stimulus, thereby prompting greater item-specific processing. Additionally, this allocated attention also results in increased processing of the relationship between the cue and target of a given word pair. Unfortunately, due to the limited nature of cognitive resources, this heightened level of individual processing inhibits relational processing between target words from different pairs. Due to this inhibition, the relational similarities between targets are far less salient in this condition. As a result, target words were more difficult to recall during testing in this condition due to the overemphasis of encoding for item-specific information as opposed to relational information between word pairs.

Although the negative testing effect was a predicted outcome of Experiment 4, comparing these results with an earlier experiment in Peterson's (2011) study yielded an unexpected result. In Experiment 3, the negative generation effect was being examined. The control condition of this experiment was a single-presentation condition, wherein subjects read a categorically organized list of rhyming cue-target word pairs once before recall testing. The unpredicted finding was that this single-presentation condition produced higher performance results than did the restudy

condition of Experiment 4, indicating the presence of a negative repetition effect.

Possible explanations At this point we will examine the two prominent explanations for these aforementioned negative effects on memory. The first of these is the notion that the negative testing and negative repetition effects are a result of how the presented information is encoded by the participant. We will begin with the negative repetition effect. In Experiment 4 of Peterson's dissertation, the restudy group was presented with the same list of rhyming word pairs twice. During the first presentation, these word pairs were presented randomly; the following presentation sorted the word pairs into categories based upon the target word. The single-presentation group in the accompanying Experiment 3 (analyzed post-hoc as a control) that was compared to this condition was presented with the word pairs in categorical groups, but did not have a prior random presentation of the items.

In addition to the item-specific versus relational account described earlier, another explanation for the reason the Experiment 4 restudy group did worse is the principle of negative transfer. Negative transfer is the concept that ineffective encoding strategies may be transferred from one list presentation to another, thereby reducing memory performance. To provide a comparison, the occurrence of negative transfer in memory encoding situations is analogous to functional fixedness in creative problem-solving (wherein suboptimal problem-solving strategies persist from one situation to the next).

To apply this concept as an explanation, one might reasonably propose that the cause for poorer recall in the restudy group is the greater amount of attentiveness subjects gave to the within-pair similarity (namely, the fact that the word pairs rhymed). Due to the fact that the first presentation was pseudo-randomized, it is unlikely that participants would have noticed the relational properties between the target words because there was no organized grouping of items. Because this between-pair categorization was unlikely to be noticed initially, the subjects might have been biased to notice only the within-pair rhyming similarity in the subsequent presentation. In contrast, the group that was only presented the items once (but in categorical groupings) might have been more likely to notice the between-pair relational information. By encoding the target words within the context of

meaningful categories, it is reasonable to assume that these subjects would gain an advantage during the free recall phase. Although repetition and the spacing effect would normally create higher recall of targets at test, such benefits were not enough to surmount the deficit caused by encoding the target items as unrelated.

There is, however, the possibility that encoding is not solely responsible for these observed discrepancies between groups. It may be the case that the observed negative repetition and negative testing effects are retrieval-based phenomena. Simply put, this notion would assert that during the final free recall test, participants in the "restudy" (repetition) and "retrieval" (testing) conditions of Experiment 4 were less able to actively recall the categories observed earlier. Due to this deficiency, they would be unable to cue themselves to enhance recall ability. However, the condition in the corresponding Experiment 3 that was only presented with the word pair stimuli once and in categorical order (the single-presentation condition) would have a higher likelihood of retrieving the category information upon testing, and as such would be able to provide self-cues to reduce the difficulty of recalling the targets during the final recall task. This alternative was not explored in Peterson's dissertation, but it remains a possible explanation for what may have been prompting these observed negative effects.

The current study Although Peterson's observations offer a compelling and unique perspective on the nature of repetition and testing effects, they are not conclusive. Peterson (2011) does not deny this, and provides a detailed account of limitations that are apparent in his study. First among these is the fact that the explanation for the negative repetition effect is entirely post-hoc. Having not set out to test this phenomenon – indeed, he was surprised himself that it occurred – he simply analyzed the data after the unexpected trend was noted. Furthermore, he again notes that the analysis involved a cross-experimental comparison. Although the main effect of the two experiments was found to be significant and the populations utilized were quite similar, there was no random assignment of subjects to conditions. Peterson believes that accounting for these two points would make this finding far more compelling (*see* Peterson & Mulligan, in-press).

With respect to the negative testing effect, Peterson's findings were more convincing, as a majority of his experiments were successful in producing the effect. Nevertheless, this study is only scratching the surface of the potential implications a negative testing effect might represent. Although there is a considerable amount of literature on the testing effect, Peterson laments that precious little is known about the process itself. As such, he feels that further emphasis should be placed on determining the mechanisms underlying the testing effect to determine exactly *why* it is that tests facilitate improvement in (or in this case, inhibition of) memory.

The current study was designed as a means of addressing these issues through the course of two experiments. In addition, the design of both experiments is also intended to determine the source of the negative testing and negative repetition effects – in other words, whether an encoding or retrieval explanation can more effectively account for the occurrence of either effect. In order to achieve this, the structure of both experiments will be based upon Peterson's (2011) dissertation as well as previous studies (Burns, 1990; Karpicke & Zaromb, 2010; Roediger & Karpicke, 2006).

For Experiment One, there will be a single-presentation control condition which will require the subject to read over a list of rhyming word pairs once. These word pairs will be organized by the category of the target word; for instance, "Linger – Finger" and "Harm – Arm" would be positioned next to one another since the target words fall within the same taxonomic category. A second condition (the "restudy" condition) will be allocated to observe how the effects of repetition compare to the control group. This condition will be presented with the same list of word pairs through two phases – first in pseudo-random ordering and then in categorical ordering. The final condition (the "retrieval" condition) will gauge the results of the direct testing effect in comparison to the restudy group. The first phase here will be identical to the restudy condition's first phase, but during the second phase participants will be asked to generate the appropriate target word when given only the accompanying cue word. For instance, if the word pair "beg – leg" was in the first phase, participants would be presented with the stimulus "beg – ___" and asked to fill the blank accordingly (importantly, the retrieval group will be given the correct answer at the end of each stimulus presentation).

Finally, Experiment One will also feature a category-cued recall test as the final phase for all conditions (as opposed to the original free recall test). By cuing all groups identically with the categories in which the earlier target words belong, the availability of the category names as retrieval cues will be equalized for all conditions. After testing, the performance of each condition will be assessed and compared. Specifically, the single-presentation and restudy conditions will be compared in order to assess the effect of repetition, and the restudy and retrieval conditions will be compared to determine the effect of testing.

The reason the comparisons are made in this way (as opposed to having both the restudy and retrieval conditions compared with the single-presentation condition) is to isolate the structural elements that vary between each condition. In other words, there are so many similarities between the conditions that it is necessary to determine which alteration is responsible for the observed outcome. By comparing the single-presentation and restudy conditions, the only difference between the groups is the inclusion of a pseudo-random word pair presentation in the latter. By comparing the restudy and retrieval conditions, the only difference is how the organized list of phase two is presented (i.e. keeping the target words initially blank so the participant has to recall the word from memory). If the retrieval condition was compared to the single-presentation condition, it would be unclear whether the results were due to the element of repetition (also found in the restudy group) or the testing element of phase two. Rather, the retrieval condition will be compared with the restudy group to determine whether the inclusion of a testing element compounds the negative effect with the repetition already inherent in both conditions.

If either (or both) of the negative effects are still present, this will provide evidence against the reliance of retrieval cues in mediating the effect(s). However, if a given negative effect does not appear after the category-cued recall test, this would suggest that retrieval plays a larger role in the facilitation of the given effect than was originally presumed.

The design of Experiment Two will be similar in nature to the first experiment with a few notable changes. First, the restudy condition will feature the categorical ordering of cue-target word pairs *before* the pseudo-randomized presentation, essentially interchanging the first

two phases of the Experiment One restudy group. Since the same phase substitution cannot be made for the retrieval condition – wherein phase two relies upon prior study of word pairs – this group will be omitted from the experiment. Finally, the category-cued recall test at the end of each condition will be replaced with a free recall test.

Comparison of these two groups will once again indicate the effect of repetition under these experimental parameters. However, in this instance the compared results have different implications. For instance, if the negative repetition effect is not present and the restudy group has comparable (or perhaps superior) performance to the single-presentation condition, this would indicate a greater likelihood of an encoding phenomenon taking place. This explanation is derived from the fact that presenting the categorically organized list of word pairs in phase one of the restudy condition should mitigate or eliminate the occurrence of negative transfer, thereby allowing for more efficient encoding of relational information in phase two (and consequently higher performance on the recall test). In contrast, if the negative repetition effect does indeed occur in this experiment, this finding would suggest that an encoding explanation is less capable of accounting for the results, suggesting the possibility of a retrieval-based explanation.

EXPERIMENT ONE

METHOD

Participants Sixty-eight participants were obtained through the Introductory Psychology subject pool at the University of North Carolina at Chapel Hill. Time spent during this experiment was allocated to each participant as a number of laboratory credits necessary for their class. There were 23 participants in the "single-presentation" condition, 23 participants in the "restudy" condition, and 22 participants in the "retrieval" condition.

Materials The critical items were a set of 36 rhyming cue-target word pairs borrowed from Peterson's (2011) dissertation (p. 56). The target words of these pairs fell into one of six different taxonomic categories, each containing six exemplars of the given category. These targets were borrowed from the category norms of Van Overschelde, Rawson, and Dunlosky (2004) – an updated and expanded list of category norms originally assembled by Battig and Montague (1969). The six categories were: "Parts of the Human Body", "Vehicles", "Kitchen Utensils", "Fruits", "Animals", and "Metals". This is also the order in which they were presented during the categorical presentation for all groups – phase one for the single-presentation condition, phase two for the restudy and retrieval conditions. To complete the word pairs, a rhyming cue-word was assigned in conjunction with each of the target words; however, the cue-words were not themselves a member of any of the six target categories (to avoid potential intrusion caused by accidental cue-word recall).

Once these word pairs were assembled, they were organized into two different lists. The first list was organized so that the target words were presented sequentially in a pseudo-random series. The purpose of pseudo-randomization of the word pairs (as opposed to unrestricted randomization) was to ensure that no two pairs with target words from the same category appeared in succession. This pseudo-randomized list appeared in phase one of the restudy and retrieval conditions, but did not appear in the single-presentation condition. The same word pairs were then assembled into a categorically organized version of the list, wherein the target words were grouped serially in relation to their taxonomic category. This list was introduced in phase one of the single-presentation condition and phase two of the restudy and retrieval condition. There was one notable distinction for the organized list in the retrieval condition; specifically, participants in this group were initially provided only the cue-word and a blank for the target word. The purpose of this was to allow subjects a period of time where they would attempt to retrieve the corresponding target from the earlier (pseudo-randomized) presentation of the word pairs.

Procedure Upon arrival, subjects were informed of their rights as research participants and then asked to sign two copies of the IRB consent form. Experimental sessions occurred with one participant and one experimenter per session (i.e. multiple trials were not conducted simultaneously).

Participants in the restudy condition underwent three experimental phases. In phase one, the participants were presented the pseudo-randomized list of rhyming word pairs

(described above). The experimenter briefly outlined the "cue-target" nature of the word pairs, and participants were instructed to read the pairs silently. They were also informed that they should attempt to learn the rhyming pairs of words for a later memory test. The word pairs were presented in the center of a computer screen in black lettering on a solid white background. Each word pair was presented individually for four seconds followed by a 500 millisecond interstimulus interval taking the form of a solid white screen. After the first phase was completed, the participant was given a math distractor task consisting of 70 arithmetic problems (which did not deviate from standard four-function mathematical notation). The participant was allotted five minutes to complete as many mathematical problems as possible without making any notes or intermediate calculations.

After the time for the distractor task elapsed, phase two of the restudy condition began. In this phase, the categorically organized version of the word pair list was utilized. Participants were reminded that later in the experiment they would be asked to remember information presented, but in this instance were specifically asked to remember the target words. Word pairs were presented more slowly in this phase (15 seconds each), and participants were asked to read each pair aloud. After this list was completed, phase three began. In this final phase, the participant was given a category-cued recall test. The participant was told that the target words came from different categories, and that category names would be presented to help the participant recall the target words. Each categorical cue was presented on the computer screen for 50 seconds for a total testing duration of 5 minutes. Participants were asked to recall as many target words that came from the category as they were able. It was made clear that each category corresponded to multiple target words, and that participants should try to recall as many targets as possible for each category. After this task was completed, the sheet with the recalled target words was collected, the participant was debriefed, and the appropriate amount of laboratory credit was assigned.

The next condition to outline is the retrieval condition, which differed in only one aspect from the restudy condition. During phase two, participants in the retrieval condition were presented with the same categorically ordered word pair list described above, but with the target word missing (i.e. the cue-word was presented in isolation). Participants were instructed to read the cue word aloud, and then say the name of the target word aloud once they recalled it. After 10 seconds, the target word was presented. If the participant incorrectly recalled the target or did not recall any target word, s/he was asked to read the target aloud. The full word pair then remained on the screen for five seconds. This was followed by a 500 millisecond interstimulus interval. While the participant was reading these word pairs aloud, the researcher was scoring the responses. A correct vocalization of the target word before it was presented was coded as a "correct response." An incorrect vocalization of the target word was coded as an "incorrect response" (the experimenter recorded the incorrect word for this response). Failure to vocalize either an incorrect or correct target word before it was presented was coded as "no response." In the rare instances in which a participant vocalized an incorrect target word followed by the correct target word, both the incorrect and correct responses were recorded (with the understanding that the participant was initially incorrect but provided a revised answer before the target word was presented). All other aspects of this condition were identical to the restudy condition.

The final condition of this experiment was the single-presentation condition. This condition was identical to the restudy condition except for the omission of phase one (and the corresponding distractor task). In other words, the only word pair list studied was the categorically organized version, and it was only studied once before the category-cued recall test. All other aspects of this condition matched the design of the restudy group.

RESULTS

In order to compare the proficiency on the category cued recall test across the three conditions, a one-way between subjects ANOVA was conducted. For all statistical tests of significance, an alpha level of .05 was utilized. The results of the analysis indicated a significant difference in the number of items correctly recalled between the single-presentation ($M = 25.00$, $SD = 5.71$), restudy ($M = 20.44$, $SD = 7.83$), and retrieval ($M = 19.36$, $SD = 4.28$) conditions; $F(2, 65) = 5.398$, $p = .007$ (Figure 1). Post hoc comparisons between individual
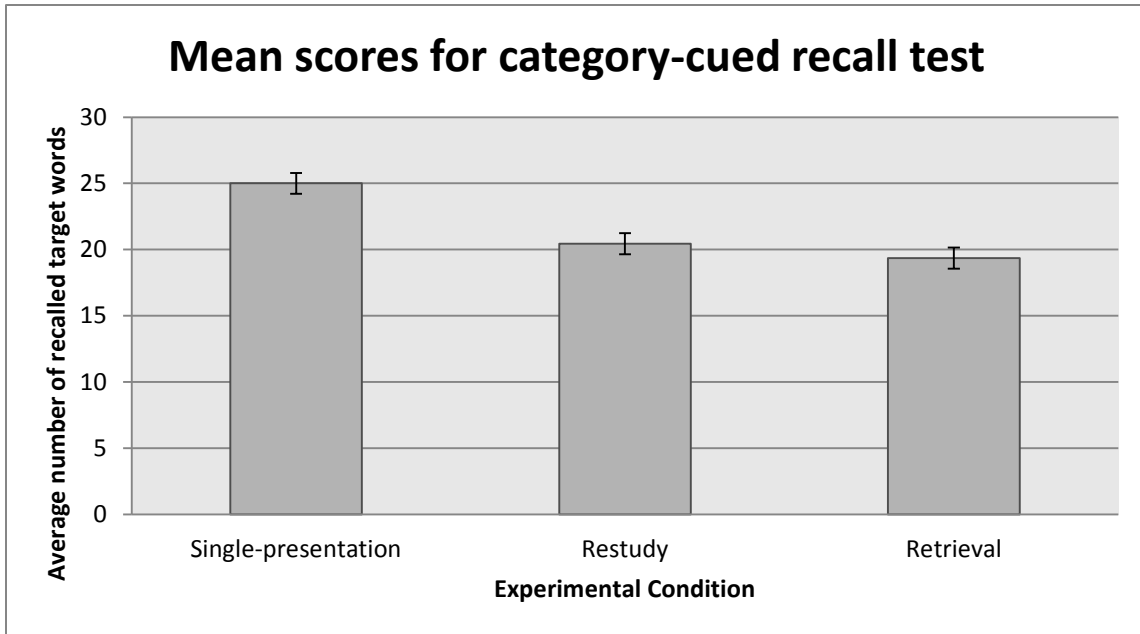
Mean scores for category-cued recall test

FIGURE 1. The compared performance on the category-cued recall test of Experiment One. There was a significant difference between the single-presentation and restudy conditions (p = .014), but not between the restudy and retrieval conditions (p = .561). Error bars represent average standard error of the means (± 0.792).

groups were assessed using Fisher's LSD, and indicated a significant difference between the single-presentation and restudy conditions (p = .014).   However, the difference between the restudy and retrieval conditions was not found to be significant (p = .561).   In other words, the single-presentation group performed significantly better than the restudy condition (suggesting the presence of a negative repetition effect), but the restudy condition failed to garner significantly higher scores than the retrieval condition (suggesting an absence of a testing effect, either negative or positive).

We also needed to determine if participants in the conditions were using different response criteria (e.g. engaging in varying levels of guessing). In order to accomplish this, a one-way ANOVA was conducted to compare the frequency of intrusions between the three conditions. Intrusions were classified as items that were reported by participants during the category cued recall test that were not actually presented during the study phase(s), but were associated with a given category. Results of the ANOVA indicated that there was not a significant difference in the mean score of intrusions among the three groups; F (2, 65) = 1.117, p = .333.   Due to this observation, we can safely claim that the occurrence of intrusions did not

impact the performance of any one condition significantly more than another.

DISCUSSION

The outcome of Experiment One revealed the occurrence of one negative effect but the absence of the other.   Namely, while the negative repetition effect was found to occur under these experimental conditions, the negative testing effect was not. Although both the restudy and retrieval conditions performed significantly worse than the single-presentation condition, this comparison was only relevant for the restudy group. In order to provide evidence for a negative testing effect, the retrieval condition would have to have performed significantly worse than the restudy group (since both of these conditions contain a form of repetition, this comparison isolates the testing group's definitive methodological variation in phase two).

These results support the use of an encoding account to explain the observed negative repetition effect.   The encoding explanation predicted that the introduction of category cues in the recall test would not prevent a negative effect from occurring. This

prediction was based upon the assumption that some conditions will facilitate greater encoding of between-pair relational information (i.e. categorical processing) than others during the study phase(s), and that this encoding variation is what accounts for the observed difference in scores. Since the presence of category cues did not eliminate the negative repetition effect, it is reasonable to infer that the difference between the single-presentation and restudy conditions arose instead during the encoding phase(s) – in other words, there was differential encoding of between-pair relational information.

In contrast, the failure to produce a negative testing effect in this experiment can be accounted for by the retrieval explanation. The retrieval explanation predicted that a recall test with category cues would prevent a negative memory effect from occurring. When participants in the restudy and retrieval conditions were supplied with the same category cues, the retrieval condition was no longer at a comparative disadvantage. This finding implies that the categorical similarities between-pairs were encoded equally in these two groups. Therefore, previous variation of performance between these two conditions – such as in Peterson's (2011) Experiment 4 – should be attributed to the reduced use of category information during recall for the retrieval condition.

EXPERIMENT TWO

METHOD

Participants Fifty-two participants were obtained through the Introductory Psychology subject pool at the University of North Carolina at Chapel Hill. Time spent during this experiment was allocated to each participant as a number of laboratory credits necessary for their class. There were 26 participants in the "single-presentation" condition, and 26 participants in the "restudy" condition.

Materials For consistency, the 36 rhyming cue-target word pairs used in this experiment were identical to those used in Experiment One. Likewise, the pseudo-randomized and categorically grouped versions of the lists used in the previous experiment were the same ones used here.

Procedure The procedure of Experiment Two matched Experiment One in most respects, but varied in a few important ways. First, this experiment did not feature a retrieval condition, and as such made no assertion as to the potential causes for the negative testing effect. Second, in the restudy condition, phase one contained the categorically grouped cue-target list and phase two contained the pseudo-randomized list, reversing the previous order of grouping presentation.

Finally, instead of using a category-cued recall test, the last phase for both the single-presentation and restudy conditions took the form of a free recall test. Participants were provided a blank sheet of paper and pen with which to record the target words presented in the earlier phase(s) of the experiment. Furthermore, they were instructed to record target words in the order that they were recalled. The testing period lasted 5 minutes.

Results For the second experiment, an independent samples t-test was used to compare the performance on the free recall test in the single-presentation and restudy conditions. Again, all statistical tests of significance used a .05 alpha level. In this instance, there was not a significant difference found in the scores for the single-presentation ($M = 21.00$, $SD = 7.56$) and restudy ($M = 23.62$, $SD = 8.59$) conditions; t(50) = -1.166, p = .249 (Figure 2). This finding indicates that switching the order of study phases in the restudy condition results in performance which does not significantly differ from the control (single-presentation) condition; in short, a negative repetition effect was not present in this experiment. On the contrary – the extent of any trending identified in the recall data is actually in the direction of a *positive* repetition effect.

Again we needed to determine if participants in the conditions were using different response criterions. Given that this experiment utilized a free recall test, there were two primary recall deviations considered. One of these was a participant's accidental recall of a cue word instead of a target word. A t-test revealed that there was not a significant difference in the occurrence of this mistake between the single-presentation ($M = .269$, $SD = .667$) and restudy ($M = .50$, $SD = 1.14$) groups; t(50) = -.891, p = .377. Intrusions were also recorded in this experiment, and were found to be equivalent for the single-presentation ($M = .462$, $SD = 1.029$)
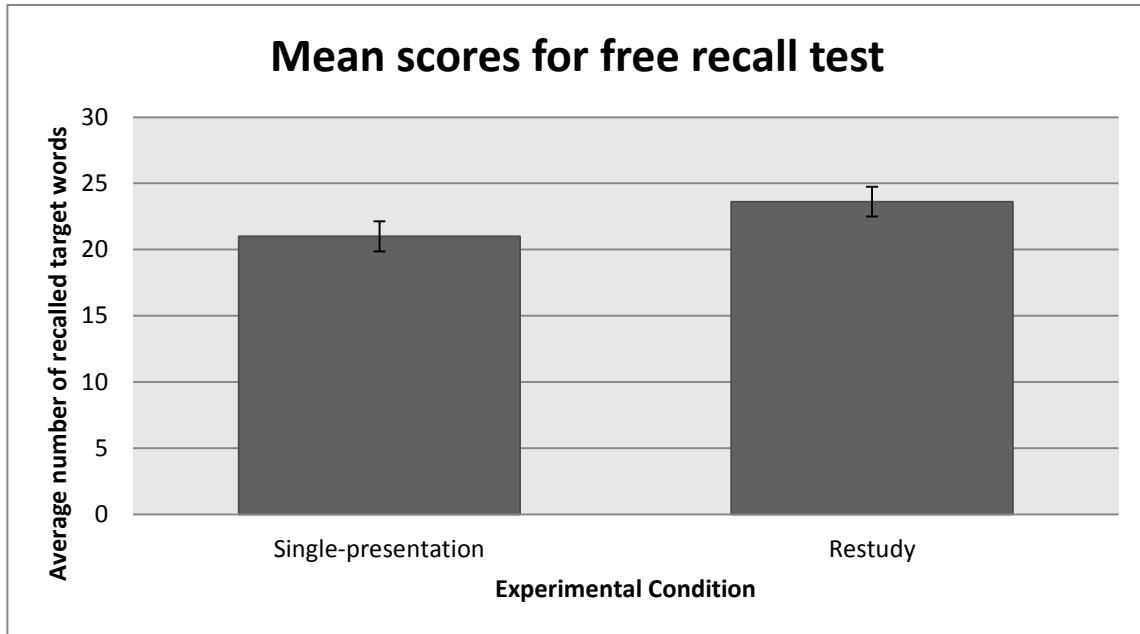
FIGURE 2. The compared performance on the free recall test of Experiment Two. The difference in scores between these groups was non-significant (p = .249). Error bars represent average standard error of the means (± 1.126).

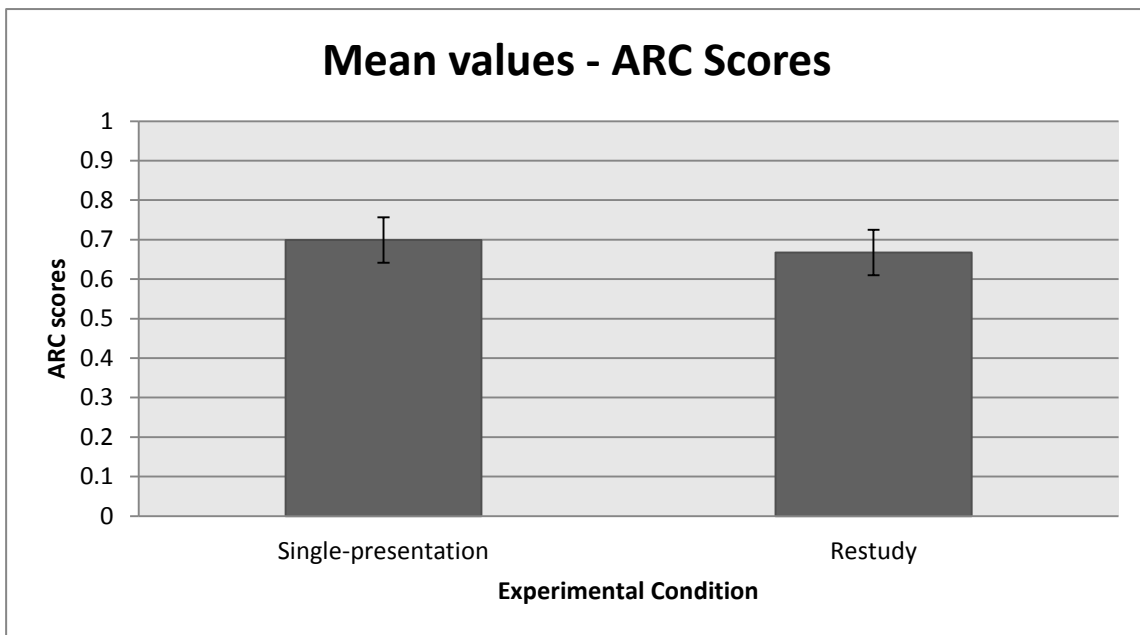FIGURE 3. The compared adjusted ratio-of-clustering (ARC) scores between the two conditions. The difference in scores between these groups was non-significant (p = .788). Error bars represent average standard error of the means (± 0.0578).

and restudy (M = .192, SD = .402) conditions; t(50) = 1.243, p = .220.

Finally, adjusted ratio-of-clustering (ARC) scores were computed by assessing the frequency with which target words of the same category were recalled in succession during the free recall test. This metric is of particular interest as it indicates either the presence or lack of categorical grouping of target items (beyond levels of chance) by subjects in a given condition. Specifically, an ARC score of 0 indicates chance-level clustering of target items, positive scores indicate above-chance frequency of clustering, and a score of 1 means that all target items were perfectly clustered (i.e. all targets were grouped in categorical succession). An independent samples t-test indicated that there was not a significant difference in ARC scores between the single-presentation (M = .70, SD = .4326) and restudy (M = .67, SD = .4084) groups; t(50) = .270, p = .788 (Figure 3). However, the scores for both conditions were positive, indicating a frequency of item clustering in both groups that surpasses levels of chance. Therefore, we can infer that while participants in the two conditions did not significantly differ in the practice of categorically grouping the target items during free recall, both conditions made use of category clustering during the recall test.

DISCUSSION

The results of Experiment Two indicated the absence of a negative repetition effect in these experimental conditions. In other words, the restudy condition did not perform worse than the single-presentation condition. On the contrary, the restudy group actually garnered a higher average score (although this difference was not statistically significant). These results are consistent with an encoding explanation of the negative repetition effect. By providing both conditions of this experiment with identical and categorically grouped initial series of word pairs, it was hypothesized that negative transfer would not occur in the restudy group (as the initial encoding was the same for both conditions). In other words, this design increased the likelihood that subjects in the restudy group would initially recognize the relational organization among the target words. Since the pseudo-randomized list was provided *after* these initial associations were made, they did not inhibit the participant's ability to group target items categorically for more efficiency in the following recall task. The

comparative performance of these two groups suggests that when initial encoding is standardized for the single-presentation and restudy conditions, the negative repetition effect dissipates.

Further evidence for the similarity of encoding between the conditions comes from a comparison of the ARC scores. Since ARC scores function as an indicator of organizational processing of items in a list, the finding that the two conditions did not significantly differ on this metric suggests that both groups utilized comparable levels of categorical processing during recall. Furthermore, the results indicated that the level of target-item clustering exceeded levels of chance for both groups. In other words, not only did both groups exhibit similar levels of categorical processing, but the grouping was too organized to be accounted for by chance. Altogether, these results reflect the reliance of participants in both conditions on between-pair relational processing.

GENERAL DISCUSSION

Taken together, these experiments successfully expanded the scope of Peterson's (2011) original study. An instance in which the negative testing effect does not occur may seem damaging at first glance, but it actually leads to a greater understanding of the conditions under which the effect surfaces by eliminating conditions under which the effect is absent. Though this result indicates a notable instance where the effect is missing, it will be the responsibility of future experiments to continue isolating instances in which the effect does take place. Furthermore, this study explored the negative repetition effect by implementing an experimental design structure, which was intended to detect such an occurrence from the start (instead of invoking cross-experimental comparisons as was done originally). The successful replication of the negative repetition effect reduces the likelihood that the original observation of its presence was anomalous, lending credence to its existence. Finally, these experiments succeeded in indicating an appropriate source for both of the negative effects – namely, the negative repetition effect was linked with the encoding explanation, and the negative testing effect was explained by the retrieval account.

Limitations Although these results broaden earlier claims concerning the negative effects of

repetition and testing, they are not without their own limitations which merit further exploration themselves. One factor which was not considered is how longer intervals of delay between word pair studying and testing may alter the expression of the aforementioned negative effects. In other words, would the same trends occur if the delay interval were a day (as opposed to the five minute distractor task utilized in this study)? Prior memory research suggests that by increasing the interval there should be a more noticeable distinction between groups utilizing effective study techniques (e.g. repeated testing) and those who are not (Roediger & Karpicke, 2006). However, would this same trend hold when experimental conditions are designed to cause one of these negative effects on memory? Answering this question is an important step to determining the relevance of these effects in an applied setting – if differences in performance do not persist over a longer period of time, then it may not be worth altering pedagogical strategies to account for this short-lived deficiency.

At this early stage in the research of these negative effects, another notable limitation is the lack of complexity in the stimuli being studied. By only using rhyming cue-target word pairs, the information being studied presumably forms fewer intricate associations between items than more complex stimuli might. In other words, a participant's processing of the association between two target words is notably simpler than the association between two *concepts*. Considering that much of what is studied in academia is frequently grouped into such abstractions, it will be important to determine whether or not the negative effects of repetition and testing directly influence only individual items of information, or larger bodies of knowledge as well.

Finally, as with most lines of research, this study may have benefited from a greater variety of represented demographics. Of these demographics, education and age seem to be the most immediately pertinent to account for. The entire sample of participants was comprised of undergraduates, and it is reasonable to suspect that college students may have more firmly grounded study habits than the general population. Due to this, the process of encoding the presented study items may vary in some key way, resulting in an altered expression of these effects – perhaps enhanced, perhaps reduced.

A similar issue may arise with children, who do not have as much practice with studying lists of information. It may be the case that the usefulness of relational associations for categorically similar information may not be as obvious for children, and therefore may not affect them in the same manner. If, however, these demographic groups were to behave in a manner similar to our undergraduate sample, then there would be greater support for the validity of applying this information both for early education and outside the realm of academia proper.

Future research One of the chief concerns for future lines of research is the construction of experiments which increasingly feature types of information which are more likely to be sequentially processed in a given setting. In doing this, the results are less confined to the purely theoretical, and modes of application become more readily apparent. Continuing with the example of educational relevance, future studies might feature a comparison between how negative effects of repetition and testing affect varying disciplines of study. For instance, it may be possible that studying vocabulary, historical facts (such as dates or event locations), and mathematical equations all prompt negative effects in study conditions similar to what was tested above. However, it may be just as likely that the nature of the material being studied moderates the potency or expression of a given negative effect. Not only would this knowledge help to shape classrooms for optimal learning, but it may also shed light on the underlying mechanisms responsible for the occurrence of the effects in the first place.

While exploring the qualitative features of the information studied in such experiments would be a meaningful step, it is also important to remember that the results of this study have indicated that the nature of the final test helps to determine whether the negative testing effect will occur. Specifically, with fewer cues the negative testing effect is found (*see* Peterson, 2011), but with useful cues explicitly provided (e.g. presentation of categories for target words) the testing effect – either positive or negative – does not occur. Considering this, future studies should utilize a variety of final memory tests in conjunction with the use of alternative forms of informational stimuli; this way, subsequent research will be able to account for experimental designs in which the negative testing effect is either expected to occur or be absent.

Studies such as these would assist in identifying how modifying qualitative aspects of

experimental design might cause either of the negative effects on memory. However, experiments which address the typical *quantity* of information retention necessary to perform well on a test would also be of value. So far the negative repetition effect and negative testing effect have exclusively been demonstrated in instances where there are only a few memorized words jotted down over the course of five minutes – in an actual testing session there tends to be greater amount and variety of information, as well as considerably more time to complete the assignment. Assuming the answers on a test could be objectively assessed as correct or incorrect (i.e. no essay or opinion based questions), it would be possible to utilize such an exam as the final recall task of an experiment somewhat resembling those conducted in the current study. For instance, historical information might be presented in such a way that emphasizes facts about individual battles rather than their context during a war (e.g. "Normandy Invasion / Eisenhower / Omaha Beach – D-Day"). Item lists such as these could be organized categorically (e.g. "World War II") or presented pseudo-randomly so that no two battles corresponding with a single war are presented in succession. When organized pseudo-randomly, the nature of these individual items should focus processing on within-stimulus characteristics which would obscure the apparent similarity between-stimuli and potentially cause negative transfer to affect subsequent presentations. Stimuli such as these are qualitatively distinct from those in earlier studies and demand a greater amount of retained information, thereby addressing both of the previously stated concerns simultaneously. Using a more natural structure such as this also helps to ensure that the negative effects of repetition and testing are less likely to be laboratory effects and will genuinely occur in more practical settings (in this instance, a history exam).

Implications Although the potential pedagogical implications will require further research to identify and validate, the theoretical impact of these experiments concerning the negative effects on memory recall is notably more immediate. Without speaking too broadly, studies such as this indicate that there may need to be some reassessment of the roles that repetition and testing play in memory retention. Whereas these strategies were once thought to uniformly and invariably improve – or at least not

reduce – one's performance on memory tasks, it now seems that this notion is either incorrect or incomplete. The recognition and encoding of relational information which meaningfully connects stimuli in a series seems to play an integral part in the determination of whether or not one's memory will improve or diminish. In short, it is not only the frequency of exposure to information or general efficacy of studying methods which affect memory performance, but also the organizational structure of the material itself.

So it seems that in some instances practice may not always ensure perfection. Though it would be unwise to suggest abstaining from repeated practice entirely (as it is effective in far more instances than it is detrimental), it no longer seems appropriate to accept its usefulness as an undisputed facet of memory research. As future studies continue to uncover variations and exceptions to the typical patterns of memory retention, we must remember that the arrangement of information being studied plays as much of a role in memory as how we endeavor to absorb that information.

References

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3, Pt.2), 1-46.

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, 41(10), 1069-1077.

Burns D. J. (1990). The generation effect: A test between single- and multifactor theories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1060-1067.

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, 79(2), 162-170.

Hunt, R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal Of Memory And Language*, *32*(4), 421-445. doi:10.1006/jmla.1993.1023

Johnstone, K. M., Ashbaugh, H., & Warfield, T. D. (2002). Effects of repeated practice and contextual-writing experiences on college students' writing skills. *Journal of Educational Psychology*, 94(2), 305-315.

Karpicke, J. D., & Zaromb, F. M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal Of Memory And*

Language, *62*(3),227-239.
doi:10.1016/j.jml.2009.11.010

Mulligan, N. W., & Lozito, J. P. (2004). Self-generation and memory. In B. H. Ross, B. H. Ross (Eds.), *The psychology of learning and motivation: Advances in research and theory, Vol 45* (pp. 175-214). San Diego, CA US: Elsevier Academic Press.

Peterson, D. (2011). *The testing effect and the item specific vs. relational account* (Doctoral Dissertation).   The University of North Carolina, Chapel Hill, NC.

Peterson, D., & Mulligan, N.W. (in press). A negative effect of repetition in episodic memory. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Roediger, H., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives On Psychological Science*, *1*(3), 181-210.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289-335.

Contributions of the Frontal and Medial Temporal Lobes to Recognition Memory Performance in Young Adults

Brielle Caserta Stark
Bryn Mawr College

Previous research with older adults suggests that the frontal lobes and the medial temporal lobes play an important role in recognition memory. The goal of this study was to investigate the impact of high/low frontal lobe and medial temporal lobe status in younger adults' recognition memory performance. Twenty-four college-aged women completed a neuropsychological battery and a recognition memory task. Behavioral and event related potential (ERP) data were collected during the recognition memory task. The results indicated that, in young adults, the medial temporal lobe showed the greatest contribution in both the old/new effect and in subsequent recollection and familiarity responses, with frontal lobe contributing minimally to either the old/new judgment or the recollection and familiarity judgment. In summary, though the frontal lobe plays a great role in the differentiation of recollect/familiar responses in older adults, we do not see this pattern in young adults. This suggests a developmental change occurring in the frontal lobe in older adulthood, where the frontal lobe assumes a bigger role in judgments succeeding the old/new judgment, either due to compensatory mechanisms or dedifferentiation.

"I enter a friend's room and see on the wall a painting. At first I have the strange, wondering consciousness, 'surely I have seen that before,' but when or how does not become clear. There only clings to picture a sort of penumbra of familiarity, - when suddenly I exclaim: "I have it, it is a copy of one of the Fra Angelicos in the Florentine Academy – I recollected it there!"

–from the *Principles of Psychology* (p. 658) by William James.

Recognition memory has long been recognized as an essential human cognitive ability that allows one to identify stimuli, such as people, places, and items, as having been previously experienced (Onyper, Zhang, & Howard, 2010). As James indicated in his passage in the Principles of Psychology, there are hypothesized to be two processes at work within recognition memory: recollection and familiarity. Recollection is the process posited to occur when a previously seen item is successfully recognized as 'old' and that this item is recognized in the presence of contextual,

or source, information. We refer to this as "details" in this study. We see this recollection occurring later in James' statement, where he implicates the source of his recognition of the painting as being from the Florentine Academy. Familiarity, on the other hand, is the successful recognition of an 'old' item without attached context or source information. This, in James' statement, is the foremost notion that he has 'experienced' the painting before, but is unsure of the specifics of this experience.

These two processes have culminated, in cognitive psychology and neuroscience literature, into a dual-process explanation of recognition memory. This dual-process account assumes that recollection and familiarity are neurologically and neuropsychologically distinct mechanisms. Investigations into the distinctness of recollection and familiarity within recognition memory include neuropsychological studies as well as scalp electroencephalography and functional magnetic resonance imaging (Onyper, Zhang, & Howard, 2010). Literature has focused on illustrating that recollection and familiarity-based responses are associated with

qualitatively different patterns of neural activity, leading to the inference that these brain regions are indeed contributing qualitatively different types of information to the ultimate recognition decision (Malmberg, 2008).

The goal of the current study is to utilize neuropsychological and EEG methodology to reproduce the finding that young adults can be differentiated into two functional groups—low and high frontal lobe and low and high medial temporal lobe functioning—in order to explore differences in old/new and recollection/familiarity behavioral judgments and neurological patterns associated with these judgments.

Application of Event Related Potentials to Recognition Memory

Event-related potential (ERPs) studies are a valuable method of investigation into the neural correlates underlying this dual process account. The ERP method extracts time-locked potentials from scalp-recorded electroencephalography records by averaging across defined conditions, and consists of a sequence of positive and negative voltage fluctuations of labeled components (Paller, Voss, & Boehm, 2007; Friedman & Johnson, 2000), which allows for indirect investigation of neural activity and, further, active brain regions, at both encoding and retrieval of information (Friedman & Johnson, 2000). Specifically, the distinct components of ERP can give us valuable information regarding the covert data: component amplitude provides an index of the extent of neural activation, such as how the component responds functionally to experimental variables; component latency, or the point in time at which the peak occurs, discloses the timing of this activation; finally, the scalp distribution component, or the pattern of voltage gradient over the scalp at any point in time, delivers information on the overall pattern of activated brain areas (Friedman & Johnson, 2000).

ERPs have a temporal resolution in the millisecond range, which allows for precise quantification of the temporal characteristics of neural activity, which is particularly important in memory research (Friedman & Johnson, 2000), though they lack the spatial resolution required to address questions regarding neural substrates, or brain regions, of specific and different processes (Rugg & Yonelinas, 2003). However, ERPs are incredibly valuable in terms of qualitative distinctions (as indicated by time

epoch, scalp distribution/topography and magnitude of signals), which can still imply distinct neural substrates (Rugg & Yonelinas, 2003). The recent advent of functional imaging has allowed for the complementation of ERP data with fMRI data, especially event-related fMRIs, coupling the spatial resolution of hemodynamic data with the temporal resolution of scalp EEG, allowing for a more complete interpretation of the processes and brain areas recruited during formation and retrieval of explicit memories (Friedman & Johnson, 2000).

Recognition memory paradigms make direct reference to previous learning, usually a series of items or words which, after a designated delay period, are tested with lists that include both these 'old' stimuli and stimuli not previously seen (new) (Friedman & Johnson, 2000). ERPs recorded about 300 to 800ms after the onset of a recognition memory test stimulus show reliable differences between these old (studied) and new (not studied) conditions (Curran, 2004), where old words elicit a larger late positive component over parietal scalp in the interval between 400-800ms than do new words (Friedman & Johnson, 2000). This effect has been referred to as the parietal old/new effect (Friedman & Johnson, 2000).

Dissociating the processes of Recollection and Familiarity: ERP, and how it is complemented by Neuroimaging Data

ERP Data. Support for the distinction of familiarity and recollection as involving different neural patterns and, perhaps, brain regions, has come from ERP investigations. In many of these studies, the Remember/Know paradigm is utilized. First introduced by Tulving in 1985, this paradigm attempts to investigate the conscious experience accompanying explicit memory retrieval, whereas participants indicate a Remember judgment only if the recognized stimuli evokes recollection of a specific source in which the stimuli was previously experienced (Tulving, 1985). On the other hand, participants would indicate a Know judgment if the recognized stimuli were thought to have been previously experienced, but that this judgment did not contain context or source information about the previous experience (Henson, Rugg, Shallice, Josephs, & Dolan, 1999). Remember judgments are thought to reflect recollective processes, and Know judgments are thought to reflect familiarity-based processes.

Encoding. It is reasoned that, if recollection and familiarity are distinct processes at retrieval, these differences might have their locus at encoding. In a study by Friedman and Trott, the ERPs associated with study items that were subsequently associated with Remember judgments (that is, a response indicating that an item has been retrieved along with its context), elicited a greater amplitude from roughly 400 to 1,100ms than those items that were subsequently unrecognized, or missed. In contrast, patterns observed to study items subsequently associated with Know judgments (thought to reflect familiarity-based retrieval) did not differ reliably from patterns elicited by items subsequently unrecognized, or missed (Friedman & Trott, 2000; Friedman & Johnson, 2000). These findings suggested that processes prompting a Remember or Know judgment might have their locus during encoding.

In an article utilizing recognition memory for pictures (where participants studied pictures of objects in two types of study blocks and subsequently made Remember-Know and source memory judgments at retrieval), investigators indicated that, when processes at encoding were investigated in reference to pictures that were successfully recognized as 'old,' a right anterior positivity at 300-450ms was observed for pictures subsequently indicated as 'remembered' and a left anterior positivity at 300-450ms was observed for pictures subsequently indicated as 'known,' which investigators believed to imply the familiarity-component (Duarte, Ranganath, Winward, Hayward, & Knight, 2004). While the onset times of these ERP patterns were similar, the topography and time course was distinguishable, which implied that recollection and familiarity could potentially present as different neural processes at encoding.

An intriguing hypothesis drawn from these data implicates various parts of the medial temporal lobe, suggesting that the perirhinal cortex subserves familiarity (sufficient for Know judgments), whereas the hippocampus proper subserves recollection (Remember judgments) (Aggleton & Brown, 1999). This hypothesis is supported by the limited investigations of the processes of Remember and Know judgments at encoding; specifically, Mangels et al. (2000) indicated that an early negativity (N340) did not differentiate ERPs elicited by items that would subsequently elicit Remember judgments from those that would elicit Know judgments, but that this data was recorded outside the hippocampus

proper (Fernandez, Effern, Grunwald, Pezer, Lehnertz, & al, 1999; Friedman & Johnson, 2000). Likewise, the hypothesis was supported in two studies indicating that slow positivity during encoding was sensitive to recollective processes in which recordings were taken from within the hippocampus proper (Fernandez, Effern, Grunwald, Pezer, Lehnertz, & al, 1999; Friedman & Johnson, 2000; Friedman & Trott, 2000).

The small amount of data revolving around the differentiation of recollective- and familiarity-based processes at encoding does not allow one to reach a definitive conclusion about potential neural implications or associated brain regions.

Retrieval

Familiarity. The research revolving around familiarity-based processes at retrieval is mixed, with some research implicating an early (300-500ms), mid-frontal, negative ERP effect, usually known as the FN400 old/new effect, in familiarity (Curran, 1999; Curran, 2000; Curran, Tanaka & Weiskopf, 2002; Curran & Cleary, 2003; Curran, 2004; Curran & Dien, 2003)

The FN400 is thought to relate to familiarity for several reasons. In one of the prime arguments for its relationship to the familiarity process, the FN400 responds similarly to studied items and similar lures, such as studied words and plurality-reversed lures (Curran, 2000), studied pictures and orientation-reversed lures (Curran & Cleary, 2003), studied geometric figures and visually similar lures (Curran, Tanaka, & Weiskopf, 2002) and studied words and semantically similar lures (Nessler, Mecklinger, & Penney, 2001; Curran, 2004). During these phenomena, the FN400 component does not change, while the component hypothesized to be related to recollection, the parietal old/new component, is affected. Finally, this FN400 begins roughly 100ms earlier than the proposed parietal old/new effect implicated in recollection, which correlates with speeded responding experiments implicating familiarity-based processes as present earlier than recollective processes (Yonelinas & Jacoby, 1995; Friedman & Johnson, 2000). Thus, the FN400 is thought to reflect familiarity (Paller, Voss, & Boehm, 2007).

However, the FN400 has been questioned in its validity for assessing familiarity, whereas some authors have not replicated the finding in amnesic patients (Curran, 2004; Olichney,

Petten, Paller, Salmon, Iragui, & Kutas, 2000), suggesting that the FN400 may be more related to a novelty-detection process downstream of familiarity judgment (Tsivilis, Otten, & Rugg, 2001), or that the FN400 could measure an unknown combination of familiarity and other co-occurring memory phenomena, like conceptual priming (Paller, Voss, & Boehm, 2007).

Recollection.   A later (400-800ms), parietal, positive ERP effect is thought to be related to recollection, and is often called the parietal old/new effect (Curran, 2004). This parietal old/new effect is maximal over left parietal electrode sites, and is topographically and functionally distinct from the mid-frontal/FN400 old/new effect (Diana, Van den Boom, Yonelinas, & Ranganath, 2011; Allan, Wilding, & Rugg, 1998; Curran, 2000; Curran & Cleary, 2003; Curran & Dien, 2003; Curran, Tanaka, & Weiskopf, 2002)

As stated earlier, recollection is characterized by the retrieval of qualitatively specific source information regarding the recognized items, and therefore the observation that this proposed parietal old/new effect is affected in experiments involving studied and similar lures (see aforementioned examples) provides consistent support for a relationship to the process of recollection (Curran, 2004). This parietal old/new effect is larger when the subsequent recognition of an item is based upon "remembering" rather than "knowing," in the Remember/Know paradigm discussed earlier (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg, Schloerschedit, & Mark, 1998; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999). The parietal old/new effect has been found to be sensitive to other variables affecting recollection, like word frequency (Rugg M. D., 1990; Rugg, Cox, Doyle, & Wells, 1995), level-of-processing (Rugg, Allan, & Birch, 2000; Paller & Kutas, 1992), and words versus pseudowords (Curran, 1999; Curran, 2004). Decisions thought not to involve recollection (incorrectly categorizing old words as new, incorrectly categorizing new words as old and correctly rejecting new words) all reflect parietal effects that are dissimilar to the proposed parietal old/new effect in recollection (Johnson, Kreiter, Russo, & Zhu, 1998a; Smith & Guster, 1993; Wilding, Doyle, & Rugg, 1995; Friedman & Johnson, 2000).

In summary, a plethora of ERP studies have hypothesized regarding the neural patterns underlying familiarity and recollection.   The FN400 and parietal old/new effects have been

dissociated in research as encompassing different time epochs and scalp topographies, and having differential effects in experimental manipulations (Curran, 2004).   The general opinion of investigators supports the implication of a parietal old/new component at retrieval for recollection, though opinions are divided on the implication of an FN400 component as representing familiarity at retrieval.   Moreover, the evidence is more conclusive of this difference at retrieval than at encoding, though the cumulative evidence does inspire the view that these differences are present in both instances and reflect different brain processes possibly underlying recollection and familiarity. Taken along with recent functional neuroimaging studies and neuropsychological results, these data imply that recollection and familiarity are supported via functionally distinct neural representations (Duarte, Ranganath, Winward, Hayward, & Knight, 2004).

Support   from   neuroimaging   and neuropsychological data

The finding that recollection and familiarity result in distinct ERP components is complemented by neuropsychological and   neuroimaging data. Patients with medial temporal lobe lesions support Aggleton and Brown's (1999) hypothesis (suggesting that the perirhinal cortex subserves familiarity, whereas the hippocampus proper subserves   recollection), implicating the hippocampus proper in recollection and the surrounding regions, like the rhinal cortex, in familiarity. Recent fMRI studies have supported this distinction. An overview of fMRI studies at retrieval has implicated that different patterns of brain activity in frontal, parietal, and medial temporal cortices make-up recollection and familiarity, and suggest that recruitment of additional brain regions in frontal and medial temporal cortices occurs during recollection (Skinner & Fernandes, 2007).

Familiarity. The ERP component, FN400, as discussed earlier, has been shown via intracranial   recordings   and magnetoencephalography (MEG) to possibly originate from anterior, inferior temporal regions like the perirhinal cortex (Duzel, et al., 2003; Curran, 2004). Further, recent fMRI research has implicated an old/new difference in the perirhinal cortex thought to be related to familiarity, such that these differences were not sensitive   to   processes   thought   to   involve

recollection, such as intentional/incidental task differences and amount of contextual information received (Curran, 2004).

The role of level of confidence was found to be important in implicating brain regions associated with familiarity judgments. In a large meta-analysis of fMRI data at retrieval, studies found that frontal BA areas 45 (pars triangular of the inferior frontal gyrus) and 6 (premotor cortex and supplementary motor area) of the left hemisphere showed high agreement in increasing activation with increasing confidence, as did left BA 39 (angular gyrus) in the inferior parietal lobe. In all of the studies, medial temporal lobe activation increased with decreasing confidence, specifically in the hippocampus (Yonelinas, Otten, Shaw, & Rugg, 2005), perirhinal cortex (Montaldi, Spencer, Roberts, & Mayes, 2006) and rhinal cortex (Daselaar, Fleck, Dobbins, Madden, & Cabeza, 2006; Skinner & Fernandes, 2007).

Recollection. The ERP left parietal old/new effect is corroborated by much neuropsychological and neuroimaging evidence. An amnesic patient with seemingly isolated bilateral hippocampal damage sustained in childhood demonstrated the FN400 old/new effect, but not the parietal old/new effect, suggesting that hippocampal activity is central to the process of recollection (Rugg & Yonelinas, 2003). The debate still continues as to whether or not the hippocampal region also contributes to familiarity, as briefly indicated in the above section; however, functional imaging studies have shown that hippocampal activity is more associated with "remember" rather than "know" judgments (Eldrige, Knowlton, Frumanski, Bookheimer, & Engel, 2000), and in source recollection (Dobbins, Rice, Wagner, & Schacter, 2003; Curran, 2004). Performance on source recollection tasks is commonly used to elucidate the processes of recollection, as the foundation of recollection relies on the ability to recall source information along with successful recognition (Friedman & Johnson, 2000). In a meta-analysis of fMRI studies at retrieval, high levels of agreement were found for activation in BA 40 in the parietal lobe, with higher concordance in the left hemisphere, and high agreement was found across the 11 studies examining activity within the medial temporal lobe (Skinner & Fernandes, 2007).

Brain regions and patterns implicated in recollection and familiarity. Thus, ERP data and complementary neuroimaging and neuropsychological evidence have contributed to the understanding of brain regions and patterns of activity associated with the processes of familiarity and recollection.

The prefrontal cortex has come under close scrutiny, as it has been realized that this area plays a critical role in both recollection- and familiarity-based processes, but that recollection probably involves additional prefrontal activity (Skinner & Fernandes, 2007). This conclusion makes considerable sense, as many studies have implicated the anterior prefrontal cortex in the retrieval of source information (Cansino, Maquet, Dolan, & Rugg, 2002; Kahn, Davachi, & Wagner, 2004; Rugg, Fletcher, Firth, Frackoqiak, & Dolan, 1996; Skinner & Fernandes, 2007). rTMS (repetitive transcranial magnetic stimulation) over dorsolateral prefrontal cortices (DLPFC) at encoding was found to significantly affect both recollection and familiarity, which is likewise corroborated by the neuroanatomy of this system, such that the DLPFC receives direct projections from the entorinal/perirhinal cortex and hippocampus (Turriziani, Smirni, Oliveri, Semenza, & Cipolotti, 2010). Aforementioned evidence suggests that familiarity and recollection, respectively, involve these brain areas, therefore upholding the finding that rTMS at encoding affects both processes. Neuropsychological analyses, such as the remember/know, receiver operation characteristic and source recognition paradigms, have also indicated that prefrontal cortex damage can impair both recollection and familiarity (Duarte, Ranganath, & Knight, 2005; Farovik, Dupont, Arce, & Eichenbaum, 2008; MacPherson, Bozzali, Cipolotti, Dolan, Rees, & Shallice, 2008; Yonelinas, Aly, Wang, & Koen, 2010). However, the precise contribution of the frontal cortex, specifically prefrontal cortex, to the processes of recollection and familiarity remains unclear.

The parietal cortex has also been implicated in the differentiation of recollection and familiarity, especially due to the ERP finding of the maximal left parietal old/new effect in recollection. In general, multiple studies have found that the parietal lobe shows greater activation for hit than correct rejection responses (Kahn, Davachi, & Wagner, 2004; Konishi, Wheeler, Donaldson, & Buckner, 2000; Wheeler & Buckner, 2004; Skinner & Fernandes, 2007). More specifically, we learn, from the same fMRI meta-analysis mentioned earlier, that both recollection and familiarity activate precuneus

regions of the parietal lobe (BA 7), but recollection also activates the inferior parietal lobe (BAs 40 and 39) (Skinner & Fernandes, 2007). This finding was subsequently corroborated by another meta-analysis, mentioning that the left inferior lateral parietal cortex (BA 39/40) has been consistently linked to recollection (Vilberg & Rugg, 2007).

The medial temporal lobe has been discussed extensively throughout this section, as research is drawn as to whether or not the medial temporal lobe supports only recollection, or both recollection and familiarity. Some research has indicated that the level of confidence of a judgment reflects uniquely within the medial temporal lobe, such that familiarity-based responding is associated with a decrease in activity in some parts of the medial temporal lobe, such as the hippocampus (Yonelinas, Otten, Shaw, & Rugg, 2005), perirhinal cortex (Montaldi, Spencer, Roberts, & Mayes, 2006) and rhinal cortex (Daselaar, Fleck, Dobbins, Madden, & Cabeza, 2006; Skinner & Fernandes, 2007). In other studies, familiarity is generally associated with the surrounding entorhinal and perirhinal volume of the medial temporal lobe (Yonelinas, Aly, Wang, & Koen, 2010). Recollection is thought to be related to the hippocampus proper (Aggleton & Brown, 1999); this is corroborated using different paradigms, such as the Remember/Know, source memory and ROC paradigms, whereas 16 of 19 studies showed that the hippocampus was involved in recollection (Yonelinas, Aly, Wang, & Koen, 2010).

Contributions of the Frontal Lobe and Medial Temporal Lobe and the Frontal Lobe Hypothesis of Cognitive Aging

We have discussed data from ERP and neuroimaging studies implicating specific brain regions and patterns of activation in the processes of recollection and familiarity. The ability to tap into these areas via neuropsychological investigation opens the doors to utilizing ERP to investigate the varied functional contribution of areas such as the frontal lobe and medial temporal lobes to judgments of recollection and familiarity, and to possibly obtain a double dissociation between brain areas utilized during tasks tapping into these processes.

Glisky et al. administered several neuropsychological tests to an elderly population in an attempt to create two independent factors measuring frontal lobe and medial temporal lobe function (Glisky, Polster, & Routhieaux, 1995). The highest loading tests on the frontal lobe factor were the Mental Arithmetic Test (consisting of 14 time-limited questions requiring mental arithmetic) and the Mental Control Test (consisting of several prompts ranging from easy to difficult, reaching from naming the months of the year to alternating counting by sixes whilst reciting the days of the week), both from the Weschler Memory Scale-Revised. The highest loading tests on the medial temporal lobe factor were the Logical Memory Test (involving the recitation of two stories, with the second story being presented twice, and requiring the verbatim response of each story from the participant), Verbal Paired Associates (involving a list of word pairs and a subsequent prompt to elicit the correct word to complete the pair) and Visual Paired Associates (involving a shape/image paired with a color and a subsequent prompt to match the shape/image with the correct color), all subtests of the Weschler Memory Scale. A long-delay cued recall subtest from the California Verbal Learning Test was also included (Glisky, Polster, & Routhieaux, 1995). The test performances of 48 older adults, between the ages of 65 and 87, allowed for subsequent classification of each individual into high or low frontal lobe functioning and high or low medial temporal lobe functioning groups.

Much research suggests that frontal lobe functioning in older adults has been shown to mediate age-related deficits in item and source memory (Glisky, Polster, & Routhieaux, 1995), in that a decreased involvement of the frontal lobe in aging populations causes poorer recollection though intact familiarity. This theory of cognitive aging is broadly known as the frontal lobe hypothesis of aging. The frontal lobe hypothesis predicts that functions principally dependent on frontal regions would decline in cognitive aging (for example, recollection), while functions fundamentally independent of frontal lobes would continue comparatively spared. The hypothesis further predicts that age-related brain change would selectively impact frontal regions. However, evidence suggesting otherwise is illustrated well by Glisky and Kong, who found that neither college freshmen nor a group of more educated young adults differed significantly from older adults on the cumulative frontal lobe factor, either in mean level of performance or in variability (Glisky & Kong, 2008). This result leads one to question the

validity of the frontal lobe hypothesis of aging's main point, which presumes that the frontal lobe is as heavily implicated in recollection in young adults as it is in older adults.

There is mounting evidence to support the view that not all cognitive functions diminish at the same rate, and that there are marked individual differences playing a part in the decline rate (Glisky, Polster, & Routhieaux, 1995). Research investigating the frontal lobe hypothesis of cognitive aging has largely assumed that all younger adults are high in frontal lobe functioning and that differentiation in level of functioning only occurs with aging. Thus, if younger adults show differentiation in level of frontal lobe functioning, it is likely that this level will be associated with variance in performance on measures of memory, in that younger adults lower in frontal lobe functioning may perform significantly worse on cognitive tasks compared to high functioning younger adults, and that this poor performance due to individual difference may show exacerbated differences with age. This finding is highly probable, as Glisky and Kong indicated that, among normal young people, the variability between the two neurocognitive domains (frontal lobe and medial temporal lobe) was sufficient for differentiation into two distinct (low versus high functioning) categories (Glisky & Kong, 2008). There is also the possibility that young adults may show completely different functional usage of the frontal lobe in recognition memory, where the frontal lobe may not play as big of a role in recollection young adults as it does in older adults.

Present Study

All evidence taken together, it may well be the case that frontal lobe function is a product contingent on individual and cognitive age differences. The present study aims reproduce the finding that a group of young adults can be differentiated into two functional groups—low and high frontal lobe and low and high medial temporal lobe functioning—as first demonstrated by the neuropsychological assessment of Glisky and Kong, in order to be the first study to test the hypothesis that differences (behavioral and neurological) in old/new and recollection/familiarity success in recognition memory tasks exist in populations other than those who are experiencing cognitive aging, and to explore if young adults show the same functional patterns (that is, determining if the judgments of old/new and recollection/familiarity implicate similar brain regions and time epochs) as we see in older adults.

Behavioral Hypotheses H$_1$: Young adults will be reliably differentiated into high and low functioning groups based on their performance on the frontal lobe and the medial temporal lobe measures in the Glisky battery.

H$_2$: High and low frontal lobe and high and low medial temporal lobe young adult groups will differ in terms of overall memory performance.

H$_3$: High and low frontal lobe and high and low medial temporal lobe groups will differ in terms of their recognition memory judgments both in terms of the number of recollection and familiar responses and their newness judgments.

Neurological (ERP) Hypotheses H$_1$: The old/new effect will be replicated in this population of young adults, resembling the old/new effect seen maximally over parietal electrodes in the epoch 300-800ms.

H$_2$: Following illustration of an old/new effect, it is hypothesized that the ERP patterns of words judged as "old" (hits) will be different in the high and low frontal lobe and high and low medial temporal lobe group.

H$_3$: ERP pattern differences will differ for recollected versus familiar judgments; it is predicted that the high medial temporal lobe group will show a more positive recollective ERP component than the low medial temporal lobe group and frontal lobe groups during recollection, if indeed the frontal lobe does not play as significant of a role in recollection young adults as it does in older adults.

METHOD

Participants. Participants were recruited via email from the population of Bryn Mawr College, an all-women's liberal arts college near Philadelphia, PA. Participants (n=24) were all female, between the ages of 18-22 and currently enrolled in the college. Of the 24 participants, 17 spoke English as their first language, with the remaining speaking English for a mean of 14.71 years. Participants were compensated $30 for the experiment. Of the 24 participants that were originally recruited, 2 were removed due to problems with EEG data collection or incompletion of the neuropsychological portion of the study.

Materials.

Neuropsychological evaluation. The neuropsychological evaluation consisted of six tasks pulled from the Glisky battery (Glisky, Polster, & Routhieaux, 1995), of which three tapped frontal lobe function (Mental Arithmetic, Mental Control and Letter Fluency) and three tapped medial temporal lobe function (Logical Memory, Verbal Paired Associates and Visual Paired Associates). These tasks were chosen because they had the highest loading on the frontal lobe and medial temporal lobe factors.

Mental Arithmetic consisted of fourteen time-limited questions that require mental arithmetic to answer. Scores were based on accuracy and response time. The Mental Control task required participants to recite as quickly and accurately as possible the days of the week, the months of the year, and the numbers 1-20 forwards and backwards as well as alternating counting by sixes and reciting the days of the week. Scores were based on accuracy and response time. In the Letter Fluency task, participants are given one minute to generate words beginning with a given letter (e.g., F, A, S). In the Logical Memory task, participants heard two short stories: the first story was presented once and the second story was presented twice. After the presentation of each story, participants were asked to verbally recall as much of the story as possible. Scores were based on a standardized assessment of how accurately the participant recalls the story. On Verbal Paired Associates, participants were read a list of word pairs. Participants were then prompted with the first word of each pair and asked complete the pair with the correct recalled word. The list was read four times and recall was assessed after each presentation. Scores were based on the number of correctly completed pairs over the four trials. Finally, on Visual Paired Associates, participants are shown a series of colors and designs and must recall which design was associated with which specific color. The list is shown six times and recall is assessed after each presentation. Scores were based on the number of correctly completed pairs over the six trials.

Administration of the tests within this battery were arranged randomly for each participant, with the exception of the Visual Paired Associates II, which, by necessity, had to be either the second-to-last or last test administered due to its delayed nature. This battery took approximately half an hour to complete, and was either done directly after the memory experiment or within a week of completion of the memory experiment.

Recognition Memory Stimuli. The recognition memory task involved a study and test phase. The task utilized a 17" computer screen monitor. In the study phase, a fixation cross appeared for 1000ms, followed by a study word that also remained for 1000ms. At test, the participant was prompted for two decisions. The first decision was an old/new decision and revolved around whether or not the word presented was or was not seen on the previous study list. After making their old/new judgment, participants were further prompted to subjectively qualify their answer. If the participant indicated that the word was "new" and therefore not seen on the study list, they were then prompted to rank this answer on a continuum as "very sure new," "sure new" or "somewhat sure new." If the participant indicated that the word at test was "old" and therefore seen on the study list, they were then prompted to rank this answer on a recollection/familiarity continuum, from "weak feeling of familiarity," "strong feeling of familiarity," "few details," or "lots of details." The meaning of familiarity versus recollection was explained to the participant before any portion of this task was completed.

All indications of subjective ranking were done using assigned computer keys. The assignment of the possible responses to keys appeared on the screen with every test item, so that subjects did not have to memorize key assignments (for example, M=old, Z=new). There was also a time limit associated with each response: subjects only had 3000 ms to make each evaluation, at which point a message reading "too slow or wrong key" would appear, and transition into the next fixation and test word would occur.

Before the EEG cap was fitted, participants completed a shortened, practice version of the task to establish that the participant understood the meaning of familiar versus recollected words. In this shortened, practice version of the task, 20 words were presented at study and 40 words at test.

When the EEG cap was successfully fitted, subjects completed the full version of this recognition memory task, which involved three blocks of study/test. At study, 150 words were presented, with 300 words presented at test. There were two version of this paradigm, which were counterbalanced between participants.

Each block took roughly 16 minutes to complete. After each block, subjects were given a break that ranged anywhere from 2 to 5 minutes in length.

Electroencephalography. The memory ERP component utilized a NeuroScan Quik-Cap 32-port electrode cap, attached to NeuroAmp and recorded via Scan 4.5 software. The EEG machinery employed six facial electrodes to control for artifacts, which were placed as such: one on each mastoid, one on each temple, and one above and below the participant's left eye. These sites were prepped with an alcohol wipe and an exfoliating scrub. The facial electrodes were first filled with a water-based gel, adhered to the skin via easily removable adhesive collars, and subsequently "twirled" with a blunt-tipped needle to loosen the skin and insert a fresh column of gel. Before the cap was fitted, the electrodes were filled with QuikCells, a technology created by NuAmp, utilizing a compressed, desiccated cell placed in the electrode reservoir. When fitted on the head, these QuikCells were filled with an electrolyte, thus allowing the electrolyte to expand the cell, much like a sponge, to target the precise scalp area with little chance of bridging. The ground electrode, placed 10cm above the nasium, was the only electrode site filled with the water-based gel. Medical mesh was placed over the cap if the cap did not fit snuggly in certain places, as was often the case near the occipital electrode sites. The cap was then plugged into the NuAmp amplifier.

Data was collected on the Scan 4.5 program. Experimenters aimed for impedances in the 5.0-25.0 kOhms range. A screenshot of the impedances was taken before and after the recognition memory task completion. The ERP data was recorded continuously throughout the memory task.

Design and Procedure Participants were each designated a 2.5 hour time slot. As the participant entered the lab, they first received their monetary compensation, filled out the informed consent form, a demographics form, and were then explained the details of the recognition memory task as described in the design. When the experimenter was sure that the participant understood the recognition memory task subjective judgment portion, the participant engaged in the shortened practice version of the task, which did not involve wearing the EEG cap. After successful completion of the practice task, experimenters asked the participants if they "felt that they used all of the keys (a range of subjective judgments) at test" to evaluate that the participant understood the task.

The participants were then fitted in the EEG cap, and connected to the amplifier. Impedances were checked and fixed (sometimes requiring extra twirling of the QuikCells or additional electrolyte solution), and the data began recording before the participant began the recognition memory task. The data was collected continuously throughout the task, even throughout the break periods, as not to miss any events. After completion of the three blocks of the task, the cap was removed and participants were allowed to clean up in the nearby restroom.

Many participants opted to stay after completion of the EEG component to do the neuropsychological evaluation component. If the participant opted to finish the neuropsychological evaluation component at a later date, they were encouraged to come back to the lab within one week of the date. The neuropsychological component was completed in a separate, smaller room from the one where the participant completed the memory and EEG portion. The experimenter sat caddy-corner from the participant.

Data Analysis

Neuropsychological. The neuropsychological tests were double scored, the first scorer usually being the administrator of the examination. Each participant's raw score on the six measures was converted to a z-score. Next, a frontal lobe score was computed for each participant by averaging the participant's z-scores on Mental Arithmetic, Mental Control, and Letter Fluency. Similarly, a medial temporal lobe score was computed for each participant by averaging the participant's z-scores on Logical Memory, Verbal Paired Associates and Visual Paired Associates. Lastly, younger adults were classified as low or high functioning based on their z-score for each of the two factors. Specifically, individuals who had an average z-score below 0 in either category were designated as 'low functioning' and individuals with an average z-score above 0 were designated as 'high functioning' in that category.

EEG Data. The EEG was digitally filtered and cleaned. First, an infinite impulse response (IIR)

filter was utilized to bandpass the data, providing a cut-off of a high of 20 Hz and a low of 1 Hz within the data in order to eliminate aberrant electrical artifacts. Eye blinks were corrected using the ocular artifact correction via a designated EOG channel and regression. In this system, well-defined blinks were identified as 'prototypes,' and designated as an artifact at the peak or trough of the signal. For each file, at least 40 of these blink-representative artifacts were identified, in order to provide sufficient power for the regression coefficient to be applied. This regression coefficient was then generalized to the entire time series by utilizing the examples of the selected prototypes. Further, a specialized filter was employed that utilized manually separated segments of 'bad' data (that is, horizontal data often accompanying blinks due to the necessity to look across and down the computer screen and response keys) and clean, artifact-free data. At least 30 horizontal segments were identified in each file, and at least 15 clean segments were identified to provide a slate with which to filter the horizontal data. EMSE's comparison of horizontal segments and clean data provided a logarithmic ratio with which to estimate how many components to remove according to the linear pattern of data. Components removed tended to be one in most cases, though there were several time series that required two components to be removed.

In few cases, channels of interest presented with higher frequency than desired, even after the above procedures had been completed. In this case, a spatial interpolation filter was applied, utilizing a probe to extrapolate data from nearby electrodes in order to estimate the frequency and amplitude of the 'bad' channel. This was done in only a few cases in channels under review.

After having completed the aforementioned steps to 'clean' the ERP data, each time series was divided into major events in order to more closely assess the data pattern of the specific channels. Events were segmented via the event review; each epoch was designated as -0.2 seconds to 1.0 seconds, in the interest of looking for the prime ERP components FN400 and the parietal old/new effect. Events were rejected on the basis of conservative estimates, whereas only events presenting with especially high frequency in one or more of the channels were rejected. In most cases, this did not consume more than one-quarter of the file. Event intervals of interest, namely Old, New,

Recollected, Familiar and New Subjective Judgments, were averaged across all channels of interest for each participant after rejections had been completed. These averages were compared to the E-prime behavioral and accuracy judgments. Again, the epoch of interest in event averaging ranged from -0.2 to 1.0 seconds, and all separated events were reviewed at 100 uV/cm.

RESULTS

Behavioral Results

Hypothesis 1 (H₁). The first analysis examined variability in younger adults' performance on the frontal and medial temporal lobe measures. That is, we aimed to illustrate if our cohort of young adults could be differentiated into two different functional groups based on low versus high function in the frontal and medial temporal lobes. The results revealed that young adult participants could be reliably differentiated by their frontal lobe and medial temporal lobe functioning. Specifically, the high frontal lobe group had a significantly higher z-score (*M*= .58, *SD*= .45) than the low frontal lobe group (*M*= -.53, *SD*= .49), *t* (21) = 5.68, *p* < .001, *d* = 2.38. A similar pattern was observed with the medial temporal group (high MTL = .53 *SD*= .35, low MTL = -.49, *SD*= .34), *t* (21) = 7.15, *p* < .001, *d* = 2.97.

Hypothesis 2 (H₂). Next, we examined whether the high/low frontal lobe and high/low medial temporal lobe groups differed in their overall recognition memory performance. An independent t-test was utilized, using correct rejections and hits from the memory paradigm as the test variables, and the group type (frontal or medial temporal lobe) as the grouping variable. For correct rejections, the low frontal lobe group (*M*=.77, *SD*=.12) was not significantly different than the high frontal lobe group (*M*=.74, *SD*=.09), *t* (21) = .71, p=.49, *d*=.30. For the medial temporal lobe group, the high group (*M*=.80, *SD*=.11) showed a trend toward significance for more correct rejections than the low group (*M*=.72, *SD*=.09), *t* (21) = -1.82, *p* = 0.08, *d* = 0.76. We next looked at hits. The high (*M*=.78, *SD*=.14) and low (*M*=.73, *SD*=.13) frontal lobe groups did not significantly differ, *t* (21) = 0.71, *p* = .35, *d* = 0.40. We saw a similar pattern for the high (*M*=.79, *SD*=.14) and low (*M*=.73, *SD*=.12) medial temporal lobe groups,

whereas there was not a significant difference in terms of hits, t (21) = -1.05, p = .31, d = 0.44.

We noted that the low versus high distinction in the medial temporal lobe group showed a trend toward significance for correct rejections. The medial temporal lobe high individuals were always better than the low individuals at correct rejections, which lead us to turn to signal detection theory to further clarify potential differences within this group.

Signal detection theory provides a way to quantify recognition memory performance that takes both memory sensitivity (d') and response bias (C) into account. The results of the independent t-tests revealed that the high medial temporal lobe group (M = 1.8, SD= .53) outperformed the low medial temporal lobe group (M = 1.27, SD= .55) on memory sensitivity (d'), t (21) = 2.35, p = .03, d = .98, but that the two groups (High M= .01, SD = .49; Low M= -.02, SD = .21) did not differ with respect to response bias (C), t (21) = .23, p = .82, d = .10.

Within the frontal lobe group, there was not a significant difference in d' (sensitivity) between the high (M=1.55, SD=.73) and low (M=1.50, SD=.47) groups, t (21) = -.20, p = .84, d = 0.09. Likewise, there was not a significant difference in C (bias) between the high (M=-.10, SD=.19) and low (M=.08, SD=.38) frontal lobe groups, t (21) = 1.44, p = .16, d = 0.65, although this p value did trend toward significance. This indicated that the high frontal lobe group tended to make more conservative judgments (less likely to say 'old' to an item that is new) while the low frontal lobe group tended to make more liberal judgments (more likely to say 'old' to an item that is new) (see Table 1). The results implicating the medial temporal lobe groups as having high sensitivity for old/new judgments is what we would expect from literature, which indicates that it is the medial temporal lobe that contributes to this first old/new judgment, rather than the frontal lobe.

Hypothesis 3 (H₃). In light of the significant difference in memory performance between the high and low medial temporal lobe groups, we examined potential differences in the subjective judgments underlying memory performance. Specifically, we examined whether high medial temporal lobe individuals were more inclined to 'recollect' a response rather than call it 'familiar,' and further, if high medial temporal lobe individuals showed higher recollection than low individuals while also showing fewer instances of familiarity, as is suggested in the literature of non-pathological aging adults.

First, it was noted that, overall, all individuals tended to make more 'recollect' than 'familiar' responses. We wanted to break this down into objective versus relative markers. To analyze this judgment dichotomy objectively, we utilized individual's 'recollect' scores (that is, of the words that they correctly designated as old, the words that they then designated as having 'remembered with detail') and divided this by the total possible number of hits, 150. Separate 2 (group: high vs. low) x 2 (judgment: recollection vs. familiarity) mixed-factorial ANOVAs were conducted to examine the effect of frontal lobe and medial temporal lobe functioning on recollection and familiarity (see Tables 2 and 3). An alpha level of p < .05 was used for all analyses conducted.

A significant main effect of judgment was found, such that more 'recollect' (M=.52, SD=.04) than 'familiar' (M=.22, SD=.02) judgment responses were made by all individuals, F (1, 21) = 30.03, p < .001, n_p^2 = .59, as was expected from a glance over the data. There was not a main effect of group, showing that the high medial temporal lobe group (M=.39, SD=0.02) was not significantly different than the low medial temporal lobe group in the type of judgments made (M=.36, SD=.02), F (1, 21) = 1.09, p = .31, n_p^2 = .05, though an interactive trend was noticed, where high medial temporal lobe individuals made more 'recollect' responses than low individuals, and high individuals made fewer familiarity judgments than low medial temporal lobe individuals (see Figure 1). Though the interaction of judgment by medial temporal lobe group was not significant, F (1, 21) = 2.58, p = .12, n_p^2 = .11, the effect size was large and, with more participants, the expectation is that the power would increase, and an interaction of judgment by medial temporal lobe group would trend toward significance, indicating that higher medial temporal lobe individuals 'recollect' significantly more than lower medial temporal lobe individuals. This is what we would expect from previous literature, and what was shown by our behavioral results.

For the frontal lobe group, the ANOVA illustrated a significant main effect of judgment, where 'recollect' responses (M=.52, SD=.04)

TABLE 1. Signal detection analyses of the differences between frontal and medial temporal lobe groups.

Group	d'	d' Sig. (p)	Effect	C	C Sig. (p)	Effect
MTL-High	1.80	**0.03**	**0.98**	0.01	0.82	0.10
MTL-Low	1.27			-0.02		
FL-High	1.55	0.84	0.09	-0.10	0.16	0.65
Fl-Low	1.50			0.08		

TABLE 2. Descriptive statistics of relative and objective judgments of recollection and familiarity in medial temporal lobe group.

	MTL	Mean	SD
Recollect	Low	0.47	0.17
	High	0.58	0.19
Familiar	Low	0.25	0.09
	High	0.19	0.11

TABLE 3. Relative and objective judgments of recollection and familiarity in medial temporal lobe group.

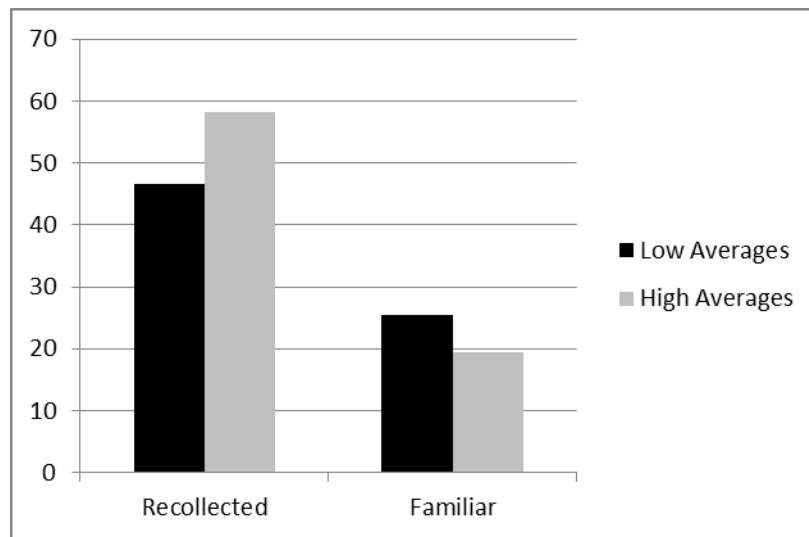| Repeated Measure Anova – Medial Temporal			
	p	Partial eta squared	Observed Power
Judgment x MTL	**0.12**	**0.11**	**0.34**
MTL	0.31	0.05	0.17
Judgment	**0.00**	**0.59**	**0.99**

FIGURE 1. Objective recollect and familiar responses grouped by medial temporal lobe. On the y-axis, the average percent correct out of 150 hits (objective marker); on the x-axis, type of objective judgment. Note the distinguishable interaction between familiarity and recollected responses in the high medial temporal lobe group.
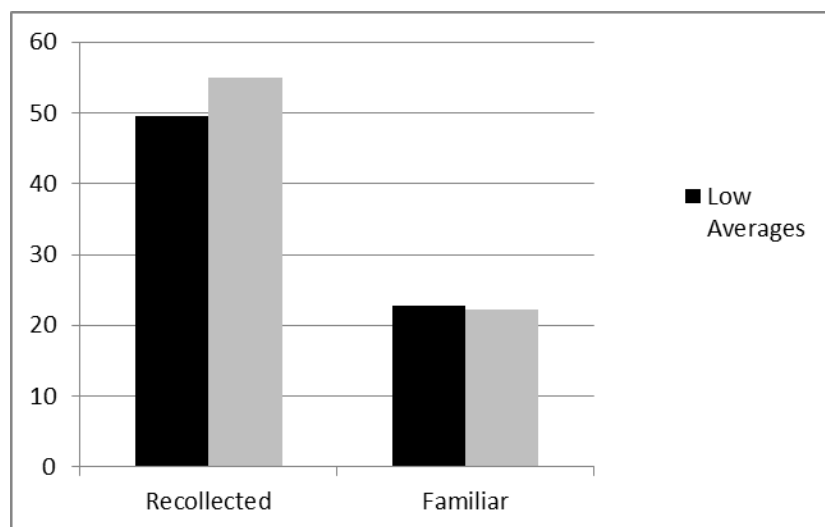

FIGURE 2. Objective recollect and familiar responses grouped by frontal lobe. On the y-axis, the average percent correct out of 150 hits (objective marker); on the x-axis, type of objective judgment. Note that there is not a distinguishable interaction across either low or high functioning groups, unlike in the medial temporal lobe groups.

were indicated significantly more than 'familiar' responses (M=.23, SD=.02), $F(1, 21) = 26.64$, $p < .001$, $n_p^2 = .56$. However, a main effect of group was not found, so that the high frontal lobe group (M=.39, SD=.02) was not significantly different in type of judgment than the low frontal lobe group (M=.36, SD=.02), $F(1, 21) = 0.83$, $p = .37$, $n_p^2 = .04$. Likewise, the interaction of judgment and frontal lobe was not significant, indicating no specific pattern in frontal lobe group response, $F(1, 21) = 0.28$, $p = .60$, $n_p^2 = .01$ (see Tables 4 and 5).

It occurred to us that being high in medial temporal lobe function would objectively contribute to the individual making more 'recollect' than 'familiar' responses, but we were

TABLE 4. Descriptive statistics of relative and objective judgments of recollection and familiarity in frontal lobe group.

	FL	**Mean**	**SD**
Recollect	Low	0.49	0.16
	High	0.55	0.21
Familiar	Low	0.23	0.16
	High	0.22	0.10

TABLE 5. Relative and objective judgments of recollection and familiarity in frontal lobe group.

Repeated Measure Anova – Frontal			
	p	Partial eta squared	Observed Power
Judgment x FL	0.60	0.01	0.08
FL	0.37	0.04	0.14
Judgment	**0.00**	**0.56**	**0.99**

TABLE 6. Descriptive statistics of relative and objective judgments of newness in the medial temporal lobe group.

	MTL	**Mean**	**SD**
Very Sure	Low	0.28	0.25
	High	0.27	0.16
Sure	Low	0.19	0.13
	High	0.27	0.11
Somewhat Sure	Low	0.23	0.14
	High	0.25	0.16

TABLE 7. Relative and objective judgments of newness in the medial temporal lobe group.

Repeated Measure Anova – Medial Temporal			
	p	Partial eta squared	Observed Power
Judgment x MTL	0.83	0.02	0.05
MTL	**0.07**	**0.15**	**0.45**
Judgment	0.57	0.02	0.09

TABLE 8. Descriptive statistics of relative and objective judgments of newness in the frontal lobe group.

	FL	**Mean**	**SD**
Very Sure	Low	0.31	0.20
	High	0.24	0.21
Sure	Low	0.25	0.13
	High	0.21	0.12
Somewhat Sure	Low	0.22	0.13
	High	0.26	0.16

TABLE 9. Relative and objective judgments of newness in the frontal lobe group

Repeated Measure ANOVA – Frontal			
	p	Partial eta squared	Observed Power
Judgment x FL	0.44	0.03	0.12
FL	0.36	0.04	0.15
Judgment	0.58	0.02	0.08

also interested in the relative judgment of the individual in reference to her own understanding of 'recollect' and 'familiar' responses. Instead of using the total number of hits possible, 150, we divided each individual's number of recollected and familiar responses by their own total hit rate, thus allowing their hit rate to function as 100% correct. In so doing, we could regard the relative judgments of the high versus low individuals in both groups.

Separate 2 (group: high vs. low) x 2 (judgment: recollection vs. familiarity) mixed-factorial ANOVAs were conducted to examine the effect of frontal lobe and medial temporal lobe functioning on this 'relative' recollection and familiarity (see Tables 4 and 5). An alpha level of $p < .05$ was used for all analyses conducted.

For the medial temporal lobe group, a main effect of judgment was seen, mirroring the objective data just described, where more 'recollect' ($M=.68$, $SD=.03$) than 'familiar' ($M=.32$, $SD=.03$) responses were indicated, $F(1, 21) = 30.05$, $p < .001$, $n_p^2 = .59$. Notably (again), the judgment by group interaction approached significance, indicating that high medial temporal lobe individuals had a tendency to make more 'recollect' judgments than low individuals, and fewer 'familiar' judgments than low individuals, $F(1, 21) = 2.37$, $p = .14$, $n_p^2 = .10$. The relative judgments in the frontal lobe group also mirrored the objective judgments of this group; results illustrated a significant main effect of judgment, where 'recollect' responses ($M=.68$, $SD=.04$) were indicated significantly more than 'familiar' responses ($M=.32$, $SD=.04$), $F(1, 21) = 26.54$, $p < .001$, $n_p^2 = .56$. Like in the objective results, the interaction of judgment and frontal lobe was not significant, indicating no specific pattern in frontal lobe group response, $F(1, 21) = 0.06$, $p = .80$, $n_p^2 = .003$.

Objective and subjective judgments of newness were also of interest to us. In the study, participants could indicate three levels of newness, ranging from 'very sure' to 'sure' and finally to 'somewhat sure.' Separate 2 (group: high vs. low) x (judgment: very sure new vs. sure new vs. somewhat sure new) mixed-factorial ANOVAs were conducted to examine the effect of frontal lobe and medial temporal lobe functioning on this newness judgment (see Tables 6-9). An alpha level of $p < .05$ was used for all analyses conducted.

For those in the medial temporal lobe group, there was not a main effect of judgment, indicating that for items indicated as 'very sure new' ($M=.28$, $SD=.04$), 'sure new' ($M=.23$, $SD=.03$) or 'somewhat sure new' ($M=.24$, $SD=.03$), there was no distinct difference in the objective choice of items correctly rejected as new, $F(2, 20) = 0.38$, $p = .67$, $n_p^2 = .02$. The main effect of the medial temporal lobe group approached significance, indicating that high individuals within the medial temporal lobe group ($M=.27$, $SD=.01$) were better, overall, at correct rejections than low individuals ($M=.23$, $SD=.01$), $F(1, 21) = 3.71$, $p = .07$, $n_p^2 = .15$ (see Tables 6 and 7). There was also no interaction between judgment and medial temporal lobe, indicating no pattern between type of judgment, $F(2, 20) = 0.05$, $p = .83$, $n_p^2 = .02$.

Looking objectively at the frontal lobe revealed no significant main effect of judgment, where judgments of new words as 'very sure new' ($M=.28$, $SD=.04$), 'sure new' ($M=.23$, $SD=.03$) and 'somewhat sure new' ($M=.24$, $SD=.03$) were not significantly different. There was also no main effect of frontal lobe group in judgments of newness, indicating that the high group ($M=.24$, $SD=.01$) and low group ($M=.26$, $SD=.01$) did not make significantly different types of judgments, $F(1, 21) = 0.88$, $p = .36$, $n_p^2 = .04$. Finally, there was no interaction of judgment with frontal lobe, $F(1, 21) = 0.63$, $p = .44$, $n_p^2 = .03$ (see Tables 8 and 9). This result echoes the literature, which does not implicate the frontal lobe in the judgment of old/new.

Finally, we wanted to look at the relative judgments of newness in both groups (thus utilizing their hit rate as 100%). In the frontal lobe, the relative judgments mirrored the objective judgments, in that no main effect of judgment was seen between 'very sure new' ($M=.36$, $SD=.05$), 'sure new' ($M=.31$, $SD=.03$) and 'somewhat sure new' ($M=.33$, $SD=.04$), $F(1, 21) = 0.08$, $p = .77$, $n_p^2 = .004$. Likewise, there not a significant interaction between judgment and frontal lobe group, $F(2, 42) = 0.47$, $p = .63$, $n_p^2 = .02$. For the medial temporal lobe group, there was not a significant main effect of judgment for the response choices: 'very sure new' ($M=.36$, $SD=.05$), 'sure new' ($M=.31$, $SD=.03$) or 'somewhat sure new' ($M=.33$, $SD=.05$). There was also not a significant interaction between judgment and medial temporal lobe group, $F(2, 42) = 0.32$, $p = .73$, $n_p^2 = .02$ (see Figures 3 and 4). These relative judgments mirrored the objective judgments.

ERP Results

Hypothesis 1 (H_1). The goal of hypothesis 1 was to replicate the old/new effect—greater
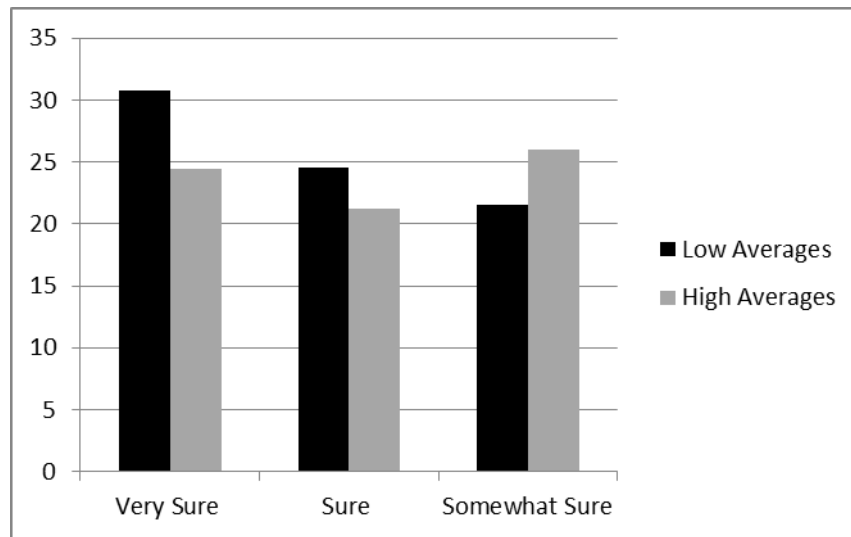
FIGURE 3. Objective newness responses grouped by frontal lobe. On the y-axis, average percent correct out of 150 hits (objective marker); on the x-axis, individual participants. Note that there is not a distinguishable interaction in either grouping condition.
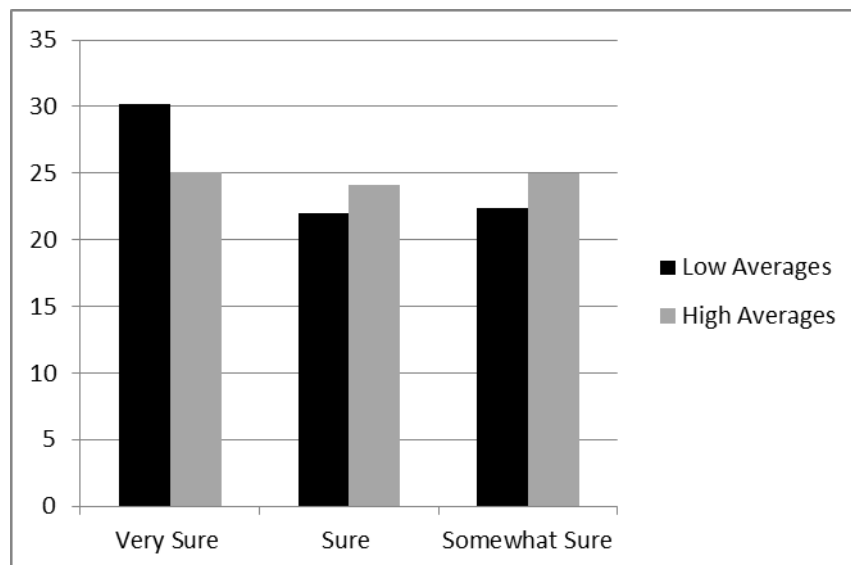

FIGURE 4. Objective newness responses grouped by medial temporal lobe. On the y-axis, average percent correct out of 150 hits (objective marker); on the x-axis, individual participants. Note that there is not a distinguishable interaction in either grouping condition.

activation for hit than correct rejection responses (Kahn, Davachi, & Wagner, 2004; Konishi, Wheeler, Donaldson, & Buckner, 2000; Wheeler & Buckner, 2004; Skinner & Fernandes, 2007)—in this population of young adults. This old/new effect has been cited as occurring maximally over left parietal electrodes in the cumulative time epoch of 300-800ms. In this study, we focused on electrodes P3 and FT7, and found

them to elicit the old/new effect, though this effect varied by functional group.

Separate 2 (group: high vs. low) x (electrode: hits and correct rejections) mixed-factorial ANOVAs were conducted to examine the effect of frontal lobe and medial temporal lobe functioning on this newness judgment. An alpha level of $p < .05$ was used for all analyses conducted. For the medial temporal lobe groups,

the main effect of old/new judgment was not significant, illustrating that the component wave of hits (*M*=.57, *SD*=.15) and of correct rejections (*M*=.62, *SD*=.19) judgments was not significantly different, *F* (1, 18) = 0.17, *p* = .69, n_p^2 = .01. The main interaction of medial temporal lobe group was also not significant, indicating that high temporal lobe individuals (*M*=.64, *SD*=.21) did not show greater positivity for hits or correct rejections as compared to low temporal lobe individuals (*M*=.62, *SD*=.19), *F* (1, 18) = 0.08, *p* = .78, n_p^2 = .08.   However, the interaction of old/new judgments and medial temporal lobe groups for ERP did approach significance, indicating that the low medial temporal lobe group showed the more replicated finding of more positive ERP components for hits than correct rejections, while the high medial temporal lobe group showed more positive ERP components for correct rejections over hits, *F* (1, 18) = 3.46, *p* = .08, n_p^2 = .16 (see Figure 5).

Hypothesis 2 (H₂). Following illustration of an old/new effect, it was hypothesized that the ERP patterns of words judged as "old" (hits) would be different in the high and low frontal lobe and high and low medial temporal lobe group.   This hypothesis investigated the notion that all correctly identified "hits" would look different across the differing functional groups; by looking for these differences, we could subsequently study the recollection and familiarity responses for each functional group.

Independent t-tests were run for all electrode channels of interest (F3, F4, FT7, FT8, T3, T4, P3 and P4) within the time epoch of 300-800 ms for events containing "hits."   For the medial temporal lobe group, several electrode channels showed greater activation (significant or approaching significance) for the low group than the high group for "hits:" **FT7** (High *M*=.05, *SD*=.47; Low *M*=.54, *SD*=.50), *t* (17) = 2.16, *p* = .05, *d* = 1.01; **FT8** (High *M*=.11, *SD*=.38; Low *M*=.52, *SD*=.50), *t* (17) = 1.98, *p* = .06, *d* = .93; **T3** (High *M*=.13, *SD*=.47; Low *M*=.46, *SD*=.30), *t* (17) = 1.87, *p* = .08, *d* = .86; **P3** (High *M*=.31, *SD*=.78; Low *M*=.86, *SD*=.56), *t* (17) = 1.83, *p* = .09, *d* = .82; and **P4** (High *M*=.39, *SD*=.58; Low *M*=.97, *SD*=.72), *t* (17) = 1.90, *p* = .08, *d* = .89 (see Table 10 and Figure 5).

Between the high and low frontal lobe groups, there was no significant difference in activation for hits across any electrodes, all *p*s greater than 0.42 and all *d*s smaller than 0.40 (see Table 11).

Hypothesis 3 (H₃). Having found a significant difference in the pattern of hits, in that there was a group difference revolving around activation levels of "hit" components at differing time epochs, the third hypothesis of this study (concerning recollection and familiarity components) could be addressed.   It was predicted that the high medial temporal lobe group would show greater activation for the 'recollective' ERP component than the low medial temporal lobe group and frontal lobe groups during recollection, if indeed the frontal lobe does not play as significant of a role in recollection young adults as it does in older adults.

Familiarity. All "hit" events marked as "familiar" (in the memory paradigm, test words either marked as remembered with 'weak feeling of familiarity' or 'strong feeling of familiarity') were considered in the early epoch of interest (300-500ms).   An independent t-test was performed comparing all electrodes at this early epoch of interest, grouping them by frontal and medial temporal lobe status.   Aging literature has suggested that the familiarity component is usually illustrated in this early epoch.   Medial temporal lobe status was significant for all electrodes in this early epoch, illustrating that the low status group showed greater activation associated with old words subsequently designated as "familiar" than the high group: **F3** (High *M*=-.91, *SD*=1.60; Low *M*=1.31, *SD*=1.93), *t* (19) = 2.79, *p* = .01, *d* = 1.25; **F4** (High *M*=-.94, *SD*=1.50; Low *M*=.79, *SD*=1.31), *t* (19) = 2.82, *p* = .01, *d* = 1.23; **FT7** (High *M*=-.83, *SD*=1.55; Low *M*=1.13, *SD*=1.80), *t* (19) = 2.62, *p* = .02, *d* = 1.17; **FT8** (High *M*=-.78, *SD*=1.37; Low *M*=.49, *SD*=1.30), *t* (19) = 2.17, *p* = .04, *d* = 0.95; **T3** (High *M*=-.56, *SD*=1.30; Low *M*=1.06, *SD*=1.79), *t* (19) = 2.29, *p* = .03, *d* = 1.04; **T4** (High *M*=-.52, *SD*=1.40; Low *M*=.97, *SD*=1.26), *t* (19) = 2.55, *p* = .02, *d* = 1.12; **P3** (High *M*=-.35, *SD*=1.62; Low *M*=2.04, *SD*=1.61), *t* (19) = 3.36, *p* < .001, *d* = 1.48;   and **P4** (High *M*=-.35, *SD*=1.45; Low *M*=2.20, *SD*=1.71), *t* (19) = 3.59, *p* < .001, *d* = 1.61 (see Table 14).   This mirrors the behavioral response described earlier, involving an interaction between the two medial temporal lobe groups, where the high medial temporal lobe group showed more overall 'recollective' responses and the low group more overall 'familiar' responses.   Frontal lobe status did not show a difference in component positivity for familiar events in any electrode channel, all *p*s
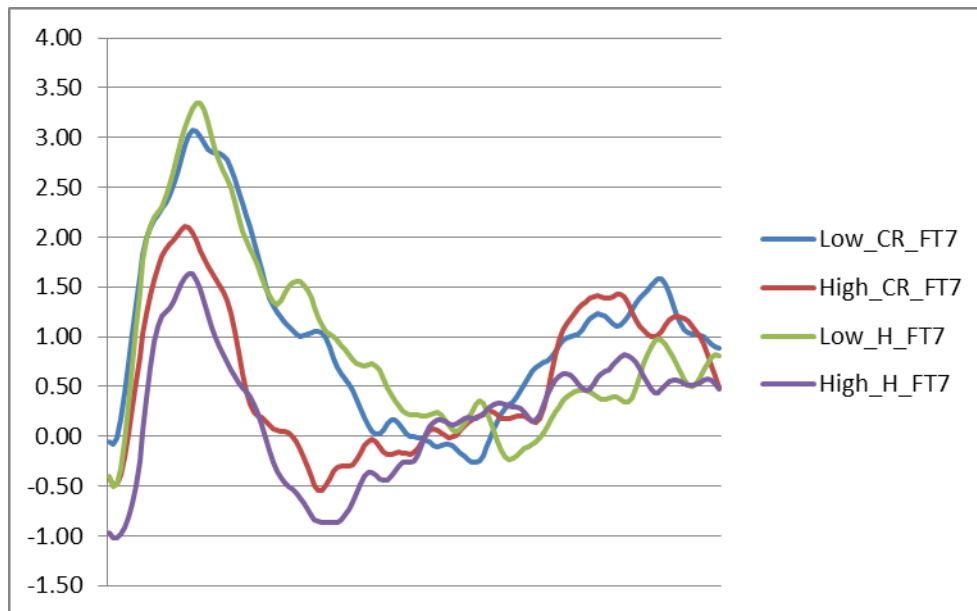
FIGURE 5. A computed ERP graph illustrating the medial temporal lobe groups for the electrode FT7 during the time epoch of 100-900 ms post-stimulus. This figure illustrates that, overall, low medial temporal lobe individuals, for both correct rejections and hits, elicit a wave with early, higher positivity (blue and green lines). This graph also illustrates that the high medial temporal lobe group (blue and purple) show the opposite pattern expected from an old/new effect, where the correct rejections are actually more positive than the hits. However, we do see the expected old/new effect in the low medial temporal lobe group (red and green), where the hit wave is more positive than the correct rejection wave.

TABLE 10. An independent t-test of "hits" comparing all electrodes at 300-800 ms epoch of interest, grouped by medial temporal lobe status.

	Low MTL Mean	High MTL Mean	t	d
F3	0.45	0.14	1.09	0.52
F4	0.40	0.11	0.96	0.46
FT7	0.54	0.05	2.16*	1.01
FT8	0.52	0.11	1.98-	0.93
T3	0.46	0.13	1.87-	0.86
T4	0.53	0.21	1.45	0.67
P3	0.86	0.31	1.83-	0.82
P4	0.98	0.39	1.90-	0.89

* = 0.05, ** = 0.0, - = trending

TABLE 11. An independent t-test of "hits" comparing all electrodes at 300-800 ms epoch of interest, grouping by frontal lobe status.

	Low FL Mean	High FL Mean	t	d
F3	0.29	0.34	-0.18	0.08
F4	0.16	0.41	-0.84	0.40
FT7	0.33	0.34	-0.03	0.02
FT8	0.29	0.41	-0.54	0.24
T3	0.30	0.34	-0.25	0.10
T4	0.41	0.39	0.61	0.04
P3	0.69	0.49	0.65	0.29
P4	0.66	0.81	-0.43	0.21

* = 0.05, ** = 0.0, - = trending

TABLE 12. An independent t-test of "recollections" comparing all electrodes at early epoch of interest, grouping by frontal and medial temporal lobe status.

	MTL-Early component (300-500ms) – Effect Size (d)	FL-Early component (300-500ms) – Effect Size (d)
F3	1.66**	0.37
F4	0.88-	0.14
FT7	1.43**	0.51
FT8	0.639	0.39
T3	1.74**	0.30
T4	0.69-	0.70
P3	1.78**	0.02
P4	1.59**	0.49

* = 0.05, ** = 0.0, - = trending

TABLE 13. An independent t-test of "recollections" comparing all electrodes at later epoch of interest, grouping by frontal and medial temporal lobe status.

	MTL-Late component (500-800ms)– Effect Size (d)	FL-Late component (500-800ms)– Effect Size (d)
F3	0.58-	0.30
F4	0.04	0.20
FT7	0.47-	0.49
FT8	0.57	0.23
T3	0.39-	0.67-
T4	0.41	0.29
P3	0.05	0.40
P4	0.32	0.13

* = 0.05, ** = 0.0, - = trending

TABLE 14. An independent t-test of "familiarity" comparing all electrodes at early epoch of interest, grouping by frontal and medial temporal lobe status.

	MTL-Early component (300-500ms) – Effect Size (d)	FL-Early component (300-500ms) – Effect Size (d)
F3	1.65**	0.19
F4	1.50**	0.01
FT7	1.51*	0.01
FT8	1.15*	0.41
T3	1.24*	0.10
T4	1.25*	0.47
P3	1.61**	0.30
P4	1.74**	0.67

* = 0.05, ** = 0.01, - = trending

TABLE 15. An independent t-test of "familiarity" comparing all electrodes at later epoch of interest, grouping by frontal and medial temporal lobe status.

	MTL-Late component (500-800ms)– Effect Size (d)	FL-Late component (500-800ms)– Effect Size (d)
F3	0.40	0.92-
F4	0.15	1.27**
FT7	0.15	0.89-
FT8	0.10	1.09**
T3	0.21	0.77-
T4	0.29	0.85*
P3	0.32	0.73-
P4	0.37	1.07*

* = 0.05, ** = 0.0, - = trending

greater than 0.32 and all ds smaller than 0.70 (see Table 14).

All "familiar" events were then considered in the later epoch of interest (500-800ms). An independent t-test was performed comparing all electrodes at this early epoch of interest, grouping them by frontal and medial temporal lobe status. Literature has suggested that the familiarity component is usually not implicated in this later epoch. In this epoch, the medial temporal lobe status did not show a difference in component positivity for familiar events for any electrode channel of interest (all ps above .52 and ds below .29). However, frontal lobe status was significant and trending across some electrodes, indicating that the high frontal lobe group showed greater activation for items subsequently labeled as familiar than the low group in this late epoch (see Table 15): **F3** (High M=.97, SD=1.44; Low M=.06, SD=.84), t (19) = -

1.83, p = .08, d = 0.80; **F4** (High M=1.37, SD=1.34; Low M=.09, SD=.49), t (19) = -3.08, p < .001, d = 1.40; **FT7** (High M=.89, SD=1.36; Low M < 0.001, SD=.95), t (19) = -1.78, p = .09, d = 0.77; **FT8** (High M=1.30, SD=1.21; Low M=.18, SD=.48), t (19) = -2.95, p = .01, d = 1.33; **T3** (High M=.76, SD=1.28; Low M=-.01, SD=.91), t (19) = -1.61, p = .12, d = .70; **T4** (High M=1.03, SD=1.30; Low M=.16, SD=.43), t (19) = -2.18, p = .04, d = 1.01; **P3** (High M=.89, SD=1.38; Low M=.16, SD=.69), t (19) = -1.59, p = .13, d = 0.71; **P4** (High M=1.21, SD=1.29; Low M=.14, SD=.67), t (19) = -2.48, p = .02, d = 1.09.

Thus, it was shown that medial temporal lobe status made a difference in component positivity towards hits marked as 'familiar' only in the early component of familiarity (which is the most implicated time epoch for this effect). Specifically, the low medial temporal lobe group showed greater component positivity for items

given a 'familiar' response than the high status group in this time period. We also saw an interesting frontal lobe status effect in the late component of familiarity (not often a time epoch implicated in the 'familiarity' decision), where the majority of electrodes indicated that high frontal lobe individuals elicited more positive components to test items subsequently labeled 'familiar' in this later epoch as compared to lower frontal lobe individuals.

Recollection. All "hit" events marked as "recollections" (in the memory paradigm, those words either marked as remembered with 'few' or 'lots of details') were considered in the early epoch of interest (300-500ms). An independent t-test was performed comparing all electrodes at this early epoch of interest, grouping them by frontal and medial temporal lobe status. Aging literature has suggested that the recollection component is not usually implicated in this early epoch.   However, the independent t-test implicated all electrode channels of interest where the medial temporal lobe status elicited significantly different 'recollective' components: **F3** (High *M*=-.75, *SD*=1.27; Low *M*=1.14, *SD*=1.24), *t* (19) = 3.42, *p* < .001, *d* = 1.51; **F4** (High *M*=-.75, *SD*=1.04; Low *M*=.14, *SD*=1.04), *t* (19) = 1.94, *p* = .07, *d* = 0.86; **FT7** (High *M*=-.52, *SD*=.95; Low *M*=.97, *SD*=1.19), *t* (19) = 3.09, *p* < .001, *d* = 1.40; **FT8** (High *M*=-.80, *SD*=.81; Low *M*=-.16, *SD*=1.01), *t* (19) = 1.58, *p* = .131, *d* = 1.05; **T3** (High *M*=-.37, *SD*=.68; Low *M*=1.00, *SD*=.79), *t* (19) = 4.17, *p* < .01, *d* = 1.86; **T4** (High *M*=-.58, *SD*=.85; Low *M*=.11, *SD*=1.13), *t* (19) = 1.54, *p* = .14, *d* = 0.70; **P3** (High *M*=-.28, *SD*=1.03; Low *M*=1.36, *SD*=.79), *t* (19) = 4.15, *p* < .001, *d* = 1.80; **P4** (High *M*=-.14, *SD*=.99; Low *M*=1.41, *SD*=.96), *t* (19) = 3.61, *p* < .001, *d* = 1.59. As is obvious, the lower medial temporal lobe group showed more component positivity for recollection responses during this time epoch (see Table 12). Frontal lobe status did not show a difference in component positivity for recollected events (all *p*s above .29, all *d*s below .48).

These "recollection" events were also considered in the late epoch of interest (500-800ms). An independent t-test was performed comparing all electrodes at this late epoch of interest, grouping them by frontal and medial temporal lobe status.   Aging literature has suggested that the recollection component is usually illustrated in this later epoch, maximal over parietal electrodes. We found that the medial temporal lobe status trended toward

significance for three electrodes, with the high group eliciting more positive components related to recollections than the low group for this time epoch (see Table 13): **F3** (High *M*=.68, *SD*=.72; Low *M*=.09, *SD*=1.02), *t* (19) = -1.49, *p* = .15, *d* = 0.68; **FT7** (High *M*=.43, *SD*=.49; Low *M*=-.02, *SD*=.83), *t* (19) = -1.46, *p* = .16, *d* = 0.68; **T3** (High *M*=.41, *SD*=.56; Low *M*=.01, *SD*=.66), *t* (19) = -1.45, *p* = .16, *d* = 0.66. Frontal lobe status did not show a difference in component positivity for recollected events except in T3, which approached significance (High *M*=-.07, *SD*=.63; Low *M*=.37, *SD*=.60), *t* (19) = 1.61, *p* = .12, *d* = 0.72. All other electrode channels for frontal lobe groups showed *p*s above .26, with all *d*s below .51.

Thus, we see that in recollection, the medial temporal lobe status makes a difference in component positivity while the frontal lobe status does not. Specifically, the low medial temporal lobe group is significantly more positive (with large effect sizes) for most of the electrodes in the early component of recollection, suggesting an earlier peak of positivity/activation for events subsequently designated as "recollected." The high group, however, shows trending effects in greater activation for recollected events in comparison to the low group in the later component of recollection (the component most usually associated with recollection).

DISCUSSION

The first analysis examined variability in younger adults' performance on the frontal and medial temporal lobe measures. The results revealed that young adult participants could be reliably differentiated by their frontal lobe and medial temporal lobe functioning. Specifically, the high frontal lobe group had a significantly higher z-score than the low frontal lobe group. A similar pattern was observed with the medial temporal group.

We examined whether the high/low frontal lobe and high/low medial temporal lobe groups differed in their overall recognition memory performance.   Results did not illustrate significant differences either between the low versus high frontal group or the low versus high medial temporal lobe group, though we noted that the medial temporal lobe high individuals were always better than the low individuals at correct rejections.

We examined potential differences in the subjective judgments underlying memory

performance. It was noted that all individuals in the study tended to make more 'recollect' than 'familiar' responses. The main difference revolved around the fact that high medial temporal lobe individuals made more 'recollect' responses than low medial temporal lobe individuals, and high individuals made fewer familiarity judgments than low medial temporal lobe individuals. There were no obvious trends between the frontal lobe groups, as was seen in the medial temporal lobe group, most likely due to lack of sensitivity for the frontal lobe group.

We also examined these potential differences in subjective judgments for newness judgments. In the frontal lobe, there was not a difference between groups as to what type of 'newness' response was made. For the medial temporal lobe group, there was no difference between the groups regarding the type of 'newness' response was made, but we did again see that the high group showed a higher overall correct rejection percentage than the low medial temporal lobe group.

Results investigating the old/new effect utilizing ERP found that, within the medial temporal lobe functional groups in electrodes P3 and FT7, those in the low medial temporal lobe condition elicited greater activation for hits than correct rejections, while the high medial temporal lobe condition showed the opposite pattern, eliciting greater activation for correct rejections than hits.

When looking just at "hits" between these groups, we see that the low medial temporal lobe group showed more positive components/greater activation early on in relation to these "old" items as compared to the high medial temporal lobe group. Frontal lobe status did not show a significant difference between "hit" components for this early time epoch.

Having found a significant difference in the pattern of hits, in that there was a group difference revolving around activation levels of "hit" components at differing time epochs within the medial temporal lobe, the third hypothesis of this study (concerning recollection and familiarity components) could be addressed. Looking at recollection in the early time epoch (300-500ms), we found that the lower medial temporal lobe group showed more component positivity for recollection responses. Frontal lobe status did not show a difference in component positivity for recollected events. In the later time epoch of recollection (500-800ms), the time period usually implicated for this effect,

we found that the high medial temporal lobe group (for some electrodes) showed more positive components related to recollections than the low group. Frontal lobe status did not show a difference in component positivity for recollected events. Looking at familiarity in the early time epoch (300-500ms), the epoch most associated with the familiarity effect, we found that the low medial temporal lobe group showed greater activation associated with old words subsequently designated as "familiar" than the high group. Frontal lobe status did not show a difference in component positivity for familiar events. In the later time epoch of familiarity (500-800ms), the medial temporal lobe status did not show a difference in component positivity. The high frontal lobe group showed greater activation for items across some electrodes for items subsequently labeled as familiar than the low group.

CONCLUSION

It is important to note that the medial temporal lobe group differentiation was due to sensitivity changes within the group, rather than bias changes within the group. This signal detection measure utilizes the conditional probability that the observer says "yes" when a stimulus is present (in our case, saying that the word is "old" when it is actually "old") and also takes into account the conditional probability when the observer says that the word is "old" when it actually is not (the false alarm rate). Thus, finding that the high and low groups were differentiated according to sensitivity level rather than bias indicated that the high group was better at filtering the 'noise' than the low group. This was supported in the behavioral data, which showed more correct 'old' (hits) and 'new' (correct rejections) responses in the high group in comparison to the low group, and in the ERP data, with illustration of an old/new effect occurring for the low medial temporal lobe group.

The ERP group differentiation brought up an unexpected trend in the high and low medial temporal lobe groups in reference to the old/new effect in P3 and FT7 and to the pattern of positivity directed toward events marked as "hits." The old/new effect usually presents as greater activation for hit than correct rejection responses (Kahn, Davachi, & Wagner, 2004; Konishi, Wheeler, Donaldson, & Buckner, 2000; Wheeler & Buckner, 2004; Skinner & Fernandes, 2007). We saw this pattern in the low medial

temporal lobe group, but surprisingly, we saw the opposite pattern (greater activation for correct rejection than hit responses) for the high group.   This is a finding worth exploring with replication. Wanting to elaborate on this finding further, we looked deeper into events designated as "hit" events—those events where a word was designated as 'old' when it was actually 'old'—and subsequent judgments concerning "hit" events, such as recollection and familiarity judgments. In the time epoch of 300-800ms, we found that, overall, the low medial temporal lobe group showed greater positivity/activation early on as compared to the high medial temporal lobe group in this cumulative "hit" analysis. This effect of the low medial temporal lobe group eliciting greater positivity/activation early on persisted when we looked more specifically into just recollection and familiar events, leading us to believe that eliciting earlier activation across all conditions was a defining factor of our low medial temporal lobe group, though the driving forces behind this are unknown and should be elaborated on further in future young adult studies utilizing this method.   One suggestion is that this memory task is much harder for the low medial temporal lobe group, thus eliciting early, higher positive activation, while the high medial temporal lobe group may have an easy time with the memory paradigm, exhibiting comparably less activation early on.

Differentiation of the groups allowed us to further explore their subsequent judgments. The data implicating high medial temporal lobe individuals in eliciting more 'recollection' responses than the low group trended toward significance, as did the data implicating fewer 'familiar' responses in the high medial temporal lobe individuals as compared to the low group. The effect sizes were moderate, and thus we can assume that, with more power, these trends would be significant. This is mirrored in the ERP results, where medial temporal lobe status was significant for all electrodes in the early epoch of familiarity (300-500ms), indicating that the low status group showed more positive ERP components associated with familiar judgments than the high group, and in the late epoch of recollection (500-800ms), with the high group eliciting more positive components related to recollection than the low group for this time epoch. These results together implicate the very interesting finding that the medial temporal lobe, at least in young adults, is heavily implicated in *both* the foremost old/new judgment and the

subsequent recollect/familiar judgments.   This finding contrasts to cognitive aging literature, which argues that the medial temporal lobe's main role is mostly old/new judgment with lesser input into recollection and familiarity.

The frontal lobe status was not shown to impact the recollect/familiar judgment, which seemed to be dictated almost exclusively by medial temporal lobe status. The only frontal lobe ERP contribution with significant differences in activation was found in the late time epoch of 'familiar' events, though the fact that we did not find statistical behavioral differences between the low and high frontal lobe groups does not allow for much elaboration on the meaning of this finding—it is unclear why the high frontal lobe group may have shown greater activation for familiar stimuli in this later time epoch. The finding that, overall, the medial temporal lobe showed the greatest contribution in both the old/new effect and in subsequent recollection and familiarity responses was highly interesting, considering the fact that cognitive aging literature suggests that recollection depends on the status of the frontal lobe: individuals with low frontal lobe function exhibit "recollect" responses far less than older individuals designated as high in frontal lobe functioning. Old age literature does corroborate our finding of the medial temporal lobe's greater contribution to the old/new effect.

Interestingly enough, it has been shown in a large study that young adults' (aged 18-23) mean frontal lobe factor was highly comparable to the mean frontal lobe factor of an older-age population (between the ages of 65 and 90) – far more comparable to this older population than a closer age-relation population of young-olds, those between the ages of 21 and 34 (Glisky and Kong, 2008). However, despite the closeness of proposed frontal lobe average scores between young and older adults, we saw a stark difference in our young adult frontal lobe contribution to recollection—or lack thereof—in this study as compared to studies implicating heavy frontal lobe contribution to recollection in older adults. In summary, the frontal lobe plays a great role in the differentiation of recollect/familiar responses in older adults. We do not see this pattern in the young adults, where the medial temporal lobe seems to play the prominent role in both the original old/new judgment and in the subsequent judgment of recollection/familiarity.   This suggests a developmental change occurring in the frontal lobe in older adulthood, where the frontal lobe

assumes a bigger role in judgments succeeding the old/new judgment, either due to compensatory mechanisms or dedifferentiation.

Limitations

During the memory paradigm, participants often shifted their eyes horizontally and downward, presumably to reassure themselves that they were pressing the key that correctly corresponded to their intended answer. This was not seen as often in the first old/new response, as these keys were pressed the most often (m=old, z=new), due to participants tending to keep their fingers stable on these with each trial. However, in subsequent judgments (recollect/familiar and newness), participants had to shift their fingers and eyes in order to choose the right key (z, x and c for newness; n, m, b and v for recollect/familiar continuum). This eye movement elicited an unexpected, large horizontal movement in the ERP data that had to be corrected. However, it was found that this correction dramatically smoothed some curves, and may have led to the inconclusiveness of the parietal old/new effect in the high medial temporal lobe group. After running the data with horizontal movements untouched, and with them subsequently retouched, it was found that the two data sets did not significantly differ and neither showed the expected parietal old/new effect in the high medial temporal lobe group.

An interesting question arises concerning the population in use in this study regarding the closeness of low and high groups within the frontal lobe group. In most conditions (and indeed, in the Glisky and Kong large study of 2008), young adults were successfully separated into statistically different low and high frontal lobe groups (Glisky and Kong, 2008). However, our study did not show this statistical difference; in fact, our frontal lobe groups looked remarkably similar to one another and only differed (though not significantly) in their types of bias. This was presumably due in large part to lack of power, but perhaps also reliant upon the population from which we drew. It is our best assumption that the population from which our participants were drawn—that of a very selective liberal arts college comprising entirely of women—attracted individuals whose frontal lobe calibers were quite similar than what we might have seen drawing from a more diverse intellectual, socioeconomic status and sex population. We did see a medial temporal lobe score difference, but again, if drawn from

another more diverse population, it is expected that this group difference would be even larger.

Finally, our study lacked sufficient power to provide statistical significance for many of the trends seen. This was undoubtedly due to the small number of participants that could be obtained and tested during the course of a busy semester. However, many of our effect sizes were very large, indicating that replication with more participants would be highly beneficial.

Future Directions

Our frontal lobe groups were not statistically different, most probably from the lack of power and the very homogenous population that we worked with. Future studies should draw from a more diverse population, to see if, first, they obtain a population whose frontal lobe scores can be successfully divided into statistically different low and high groups, and second, to see if this successful difference contributes at all to overall memory judgments (hits and correct rejections) and to subsequent judgments of familiarity/recollection. As we said, we did not find any frontal lobe contribution to either the old/new judgment or subsequent judgments, though this finding may differ with a larger population and more qualitatively/quantitatively different low and high groups.

In the future, a method that does not rely so much on checking oneself during the memory paradigm—and therefore causing unnecessary horizontal and downward eye movements—should be investigated. Thorough cleaning of the ERP data may have contributed to the surprising high medial temporal lobe parietal old/new effect seen in our study, as much of the data may have been averaged to the point where significant differences were lost.

Finally, the most broad take-away of our study revolves around determining if the developmental change in older adults' frontal lobes—from having supposedly little contribution to old/new or subsequent judgments to having large contributions to recollection as one ages— is compensatory or due to dedifferentiation of the brain's systems.

References

Aggleton, J., & Brown, M. (1999). Episodic memory, amnesia, and the hipocampal-anterior thalamic axis. *Behavioral Brain Science*, 425-489.

Allan, K., Wilding, E. L., & Rugg, M. D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Psychologica*, 231-252.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex*, 1048-1056.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia*, 771-785.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 923-938.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, 1088-1106.

Curran, T., & Cleary, A. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, 191-205.

Curran, T., & Dien, J. (2003). Differentiating amodal familiarity from modality-specific memory processes: An ERP study. *Psychophysiology*, 979-988.

Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective, & Behavioral Neuroscience*, 1-18.

Daselaar, S. M., Fleck, M. S., Dobbins, I. G., Madden, D. J., & Cabeza, R. (2006). Effects of healthy aging on hippocampal and rhinal memory functions: An event-related fMRI study. *Cerebral Cortex*, 1771-1782.

Diana, R. A., Van den Boom, W., Yonelinas, A., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, 278-286.

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia*, 318-333.

Duarte, A., Ranganath, C., & Knight, R. T. (2005). Effects of unilateral prefrontal lesions on familiarity, recollection, and source memory. *Journal of Neuroscience*, 8333-8337.

Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cognitive Brain Research*, 255-272.

Duzel, E., Habib, R., Schott, B., Schoenfeld, A., Lobaugh, N., McIntosh, A. R., et al. (2003). A multivariate spatiotemporal analysis of electromagnetic time-frequency data of recognition memory. *Neuroimage*, 185-197.

Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences of the United States of America*, 5973-5978.

Eldrige, L. L., Knowlton, B. J., Frumanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, 1149-1152.

Farovik, A., Dupont, L. M., Arce, M., & Eichenbaum, H. (2008). Medial prefrontal cortex supports recollection, but not familiarity, in the rat. *Journal of Neuroscience*, 13428-13434.

Fernandez, G., Effern, A., Grunwald, T., Pezer, N., Lehnertz, K., & al, e. (1999). Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science*, 1582-1585.

Friedman, D., & Johnson, R. J. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, 6-28.

Friedman, D., & Trott, C. (2000). An event-related potential study of encoding in young and older adults. *Neuropsychologia*, 542-557.

Glisky, E. L., & Kong, L. L. (2008). Do young and older adults rely on different processes in source memory tasks? A neuropsychological study. *Journal of Experimental Psychology*, 809-822.

Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychologia*, 229-235.

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *The Journal of Neuroscience*, 3962-3972.

Johnson, R. J., Kreiter, K., Russo, B., & Zhu, J. (1998a). A spatio-temporal analysis of recognition-related event-related brain potentials. *International Journal of Psychophysiology*, 83-104.

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuoanatomical correlates of recollection: Implications for models of

recognition memory. *The Journal of Neuroscience*, 4172-4180.

Konishi, S., Wheeler, M. E., Donaldson, D. I., & Buckner, R. L. (2000). Neural correlates of episodic retrieval success. *Neuroimage*, 276-286.

MacPherson, S. E., Bozzali, M., Cipolotti, L., Dolan, R. J., Rees, J. H., & Shallice, T. (2008). Effect of frontal lobe lesions on the recollection and familiarity components of recognition memory. *Neuropsychologia*, 3124-3132.

Malmberg, K. J. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, ??

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus*, 504-520.

Nessler, D., Mecklinger, A., & Penney, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research*, 283-301.

Olichney, J. M., Petten, C. V., Paller, K. A., Salmon, D., Iragui, V. J., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological measures of impaired and spared memory. *Brain*, 1948-1963.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology*, 341-364.

Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support of the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 375-391.

Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, 243-250.

Rugg, M. A., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Science*, 313-319.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, 367-379.

Rugg, M. D., Allan, K., & Birch, C. S. (2000). Electrophysiological evidence for the modulation of retrieval orientation by depth of study processing. *Journal of Cognitive Neuroscience*, 664-678.

Rugg, M. D., Cox, C. J., Doyle, M. C., & Wells, T. (1995). Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia*, 471-484.

Rugg, M. D., Fletcher, P. C., Firth, C. D., Frackoqiak, R. S., & Dolan, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*, 2073-2083.

Rugg, M. D., Schloerschedit, A. M., & Mark, R. E. (1998). An electrophysiological comparison of two indices of recollection. *Journal of Memory and Language*, 47-69.

Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia*, 2163-2179.

Smith, M. E., & Guster, K. (1993). Decomposition of recognition memory event-related potentials yields target, repetition, and retrieval effects. *Electroenceph Clin Neurophysiol*, 335-343.

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology & Aging*, 390-413.

Tsivilis, D., Otten, L. J., & Rugg, M. D. (2001). Context effects on the neural correlates of recognition memory: An electrophysiological study. *Neuron*, 497-505.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, 1-12.

Turriziani, P., Smirni, D., Oliveri, M., Semenza, C., & Cipolotti, L. (2010). The role of the prefrontal cortex in familiarity and recollection processes during verbal and non-verbal recognition memory: An rTMS study. *NeuroImage*, 348-357.

Vilberg, K. L., & Rugg, M. D. (2007). Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. *Neuropsychologia*, 2216-2225.

Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomical correlates of remembering and knowing. *Neuroimage*, 1337-1349.

Wilding, E. L., Doyle, M. C., & Rugg, M. D. (1995). Recognition memory with and without retrieval of context: An event-related potential study. *Neuropsychologia*, 1-25.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory & Language*, 622-643.

Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, 1178-1194.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience*, 3002-3008.

Papers submitted to the Yale Review of Undergraduate Research in Psychology are reviewed using a blind review procedure.

The lead author of each paper submitting to this journal must have conducted her or his research while an undergraduate (or equivalent) student at an accredited college or university anywhere in the world.

For more information, visit www.yale.edu/yrurp or email psych.review@yale.edu.