The impact of acute psychosocial stress on false memory

Michelle Henry

Department of Psychology

University of Cape Town

Supervisor: Kevin Thomas

Word Count:

Abstract: 235

Main Body: 10412

ABSTRACT

The current research aimed to confirm that acute psychosocial stress can lead to increased rates of false memory errors in humans. In addition, it attempted to show that false memory rates differed depending on the original stimulus type, thus extending and validating the research done by Gallo and colleagues (2004) on material specificity in false memory. Participants in an experimental group (3 males and 3 females) were exposed to a procedure designed to induce mild psychosocial stress, whereas participants in a control group (6 males and 6 females) were exposed to a period of relaxation. Salivary cortisol and subjective (self-report scales) measures were used to determine participants' stress levels. All participants completed a false memory task, entailing 3 different recognition tests, on 2 consecutive days. Although the Stress group made more false memory errors, between-group differences were not statistically significant. With regard to material specificity, in both groups false memory for words was not statistically significantly greater than those for pictures. Finally, with regard to changes from day 1 to day 2, participants in the Stress group showed a greater increase in false memory errors than did the Control group, but, again, the magnitude of change was not statistically significant. Future research efforts will include larger samples sizes as the current study's small *N* did not yield enough power for definite conclusions to be drawn.

Keywords: stress, false memory, material specificity, hippocampus, cortisol, recognition

Psychological stress, in both chronic and acute forms, is associated with a variety of cognitive, physiological and behavioural responses. This is in part due to the fact that hormones released by the body during stressful experiences regulate brain regions involved in those responses (Lupien et al., 1997; Sapolsky, Krey, & McEwen, 1986). One of the most studied effects of stress is its impact on memory, with results from both human and animal studies often indicating a negative relationship between the two (e.g., Newcomer et al., 1999). False memory is one type of memory process that has been shown to be affected by stress. There is a definite need for further research in this area, with results possibly having important implications for real-life situations such as eye-witness testimony or stressful situations such as exams. Even though artificially conducted laboratory tests may not truly reflect real-life situations, they provide a practical means of examining the impact of stress on false memory.

FALSE MEMORY

Although memory can achieve high levels of accuracy, it is susceptible to a variety of distortions and illusions (Roediger, 1996). False recognition (claiming to have previously encountered a novel word or event) is one type of memory distortion that has recently received much attention (Budson, Dodson et al., 2005).

Bartlett (1932) is usually accredited with conducting the first experimental investigation of false memories, in which subjects were asked to read an Indian folk tale and then recall it repeatedly. His results showed distortions in subjects' memories over repeated attempts to recall the story. Although many of his results have not been replicated in subsequent studies, Bartlett made a major contribution in distinguishing between reproductive memory (accurate material from memory) and reconstructive memory (filling in missing elements while remembering, with errors often occurring; Roediger & McDermott, 1995). Bartlett assumed that materials rich in meaning (such as stories and real life events) would give rise to reconstructive memory processes and therefore errors, while simplified material (such as nonsense syllables and word lists) would give rise to reproductive memory, and therefore more accurate memory. Numerous subsequent studies have shown that this is not the case, with nonsense syllables and word lists providing an adequate means of inducing false memory in the laboratory (e.g., Roediger & McDermott, 1995). Studying false memory in the laboratory has obvious advantages over real-life false memory phenomena, which are characterized by a lack of control in: a) the nature of the event, b) events occurring during the retention interval, and c) the manner in which memories are elicited (Payne, Elie, Blackwell, & Neuschatz, 1996).

Gallo and colleagues (2004) tested false memory in a laboratory experiment. Twenty-four participants (all healthy undergraduate students) were required to study a list of 288 words on a computer screen. Each word, which was printed in black font, was either followed by the same word printed in red font or a corresponding picture of that word. Some black words were followed by red words, others by pictures, and others by both red words and pictures.

At the end of the study phase, participants were given three recognition tests. Firstly a standard recognition test, followed by two criterial recollection tests (see the *Materials* section for a full description of each recognition test).

Gallo and colleagues (2004) reported a *picture superiority effect*, in which true memory for pictures was significantly higher than that for red words on the standard recognition test, and pictures hits in the picture recollection test were significantly higher than red word hits in the word recollection test. In addition, false alarms to new items in the criterial recollection tests were lower in the picture test compared to the word test.

The effect observed in the Gallo et al.(2004) study is well-known to cognitive neuropsychologists as *material specificity*. This term refers to the fact that different brain areas are involved in processing different kinds of stimuli, and that this processing of information is dependent on the type of original material used (Grady, McIntosh, Rajah, & Craik, 1998). Memory for pictures and words is material-specific in that these processes are unique to the type of material used. Both word lists (e.g., Rajaram & Roediger, 1993) and pictures (e.g., Schacter, Koutstaal, Johnson, Gross, & Angell, 1997) have successfully been used in laboratory experiments to induce false memory. A picture superiority effect is often seen with regard to memory, in which pictures and events rich in detail are more likely to be remembered (Budson, Droller, et al., 2005; Snodgrass & Vanderwart, 1980).

Numerous studies (e.g., Budson, Droller, et al., 2005; Schacter, Israel, & Racine, 1999) have reported lower false recognition for pictures compared to words, a result consistent with the general picture superiority effect. Furthermore, studies have shown that true recognition is higher for pictures compared to words (Budson, Sitarski, Daffner, & Schacter, 2002). The *distinctiveness heuristic* provides an explanation for these results: The distinctive features of pictures result in greater confidence and accuracy, whereby false recollections will fail to correspond with subjects actual recollective expectations. Reduction in false memory for pictures has been found in experimental paradigms involving both semantically related and semantically unrelated pictures (Budson et al., 2005); which is not the case for wordlists. Numerous studies have repeatedly demonstrated high false recognition rates for semantically related words, as appear; for example, in the Deese-Roediger-McDermott paradigm (DRM; Roediger & Mc Dermott, 1995) (see Appendix A for a full explanation). These high rates of false memories are most likely attributable to confusions of familiarity between the overall theme of the list and specific items (McDermott, 1996). False recognition rates for semantically-related words can be reduced, however, by pairing words with pictures during the study phase of the experimental paradigm (Gallo et al., 2004; Schacter et al., 1997).

Studies have also shown a faster response reaction time for individuals remembering pictures compared to words (Gallo, Kensinger, & Schacter, 2006). Gallo and colleagues (2006) speculate these differences in response rate might be accounted for by the additional post-retrieval monitoring processes (such as searching for additional recollective information) used when trying to remember words. Moreover, groups remembering pictures rely solely on the distinctiveness heuristic to eliminate false intrusions, resulting in faster reaction times.

Theories of false memory

There are a number of theories explaining why and when false memories occur. Some argue that false memories originate during encoding processes, while others emphasise retrieval processes. For instance, those researchers who endorse the encoding-based theories note that during encoding people must differentiate between what occurs externally and the thoughts aroused by these external events; an inability to make such differentiations might lead to false memories (Boyer, Phillips, Rousseau, & Ilivitsly, 2007). On the other hand, those researchers who endorse the retrieval-based theories note that during retrieval, strategic monitoring processes are used to determine whether the information they are remembering is accurate or not. These monitoring processes depend on a variety of factors, including presentation rate, format, modality and number of presentations, and errors during any of them might lead to false memories (Gallo & Roediger, 2002; McDermott & Watson, 2001). Importantly, Roediger and McDermott’s (1995) original DRM paper suggests false memories may be created in part during the testing phase (when participants are completing the recognition tests), with retrieval processes contributing significantly to false recall and recognition phenomena. Although theories regarding false memory abound, the review below only discusses those relevant to this study.

According to *dual process theories* of false memory, both recollection (recalling details of prior occurrence of an event) and familiarity (feeling that an event had previously occurred without recall of detailed information) contribute to our ability to discriminate between studied and non-studied items (Curran & Cleary, 2003; Gallo et al., 2004). A sense of familiarity leaves individuals with the difficult task of deciding whether they actually encountered an event or merely thought of it during the encoding process (Payne, Nadel, Allen, Thomas, & Jacobs, (2002). By this account, the semantic overlap between presented words and the non-critical lure in the DRM paradigm would leave participants with a feeling of familiarity, causing them to falsely recognise the non-presented lure. Familiarity might also lead to source monitoring errors, where people retrieve fragments of an episode but are unable to recollect how or when the information was acquired (Johnson & Raye, 1998). Deficits in source monitoring may be due to an impairment in attribution processes as well as disruption in encoding qualitative characteristics of an event (Dab, Claes, Morais, & Shallice, 1999), resulting in the construction of false memories.

As noted earlier, humans have the ability to remember pictorial materials with far greater accuracy than words (the picture superiority effect). Pictures, compared to words, have more perceptual details associated with them, which reduces the amount of source monitoring errors made and consequently lowers rates of false recognition (Budson et al., 2002; Johnson, Hashtroudi, & Lindsay, 1993). This phenomenon can be understood in terms of the distinctiveness heuristic (Schacter et al., 1999), a retrieval orientation which states that recollective expectations guide our memory decisions, with more distinctive events being easier to separate from one another during recall/recognition tasks (Gallo et al., 2004). Distinctiveness, in this context, refers to the complexity and uniqueness of the perceptual features of a stimulus. Memory monitoring processes capitalise on such uniqueness by evaluating memory for their match with the expected characteristics of a given source (Johnson & Raye, 1998).

In many recognition tests, good performance is not merely a matter of remembering whether one has *ever* seen the presented words, but rather, remembering whether one has seen those words in the *specific* experimental context (i.e., the study phase of the experiment; Payne et al., 2002). Individuals need to recruit a variety of decision-making strategies when engaged in memory-based tasks, with inconsistencies in the application of these strategies resulting in false memory production (Budson, Dodson, et al., 2005). These inconsistencies can arise as a result of cognitive aging (e.g., older adults appear to be impaired in their ability to remember the source(s) of recently acquired information, and so are more susceptible to producing false memories; Schacter et al., 1997), distraction/inattention, and stress (Johnson & Raye, 1998).

STRESSORS AND THE PHYSIOLOGICAL STRESS RESPONSE

The effects of perceiving an environmental stressor are mediated through a neuroendocrine cascade, the final result (in humans) being the secretion of cortisol, which is regarded as an objective measure of psychological and physiological stress (Kirschbaum & Hellhammer, 1992). This physiological stress response begins when the brain perceives an environmental stressor. This perception triggers the release of corticotrophin-releasing hormone (CRF) from the hypothalamus. This release in turn triggers the anterior pituitary to release adrenocorticotrophic hormone (ACTH), and this release triggers the secretion of glucocorticoids from the adrenal gland (Sapolsky et al., 1986).

The secretion of glucocorticoids from the adrenal gland protects the brain against adverse events, such as susceptibility to infectious diseases and chronic fatigue syndrome, and is essential for optimal cognitive and physiological functioning (de Kloet, Oitzl, & Joëls, 1993; Kudielka & Kirschbaum, 2005). In the brain these corticosteroids, along with other components of the stress system, co-ordinate an organism's ability to cope with environmental stressors by increasing the amount of readily available energy, increasing cardiovascular tone and altering cognition (de Kloet, Oitzl, & Joëls 1999; Sapolsky et al., 1986). Although a certain level of arousal is needed for an individual to cope with an environmental stressor, acute (Kirschbaum, Wolf, May, Wippich, & Hellhammer, 1996; Lupien, Gillin, & Hauger, 1999) or chronic (Mendl, 1999; Wolkowitz et al., 1990) elevations in cortisol levels can result in cognitive dysfunction, including memory and attention problems.

Glucocorticoids readily enter the brain and alter gene expression by binding to intracellular receptors. Corticosteroid hormone action involves binding to two intracellular glucocorticoid receptors: type 1 mineral corticoid receptors (MRs) and type 2 glucocorticoid receptors (GRs), which bind cortisol with different affinities (de Kloet et al., 1999). MRs bind naturally circulating cortisol with high affinity, whereas GRs become occupied after stress or circadian peak (Newcomer et al., 1999). MRs are involved in behavioural reactivity to novel situations, whereas GRs are involved in consolidation and storage of learned information (e.g., Kirschbaum et al., 1996). de Kloet et al. (1999) showed that activation of both types of receptor is a prerequisite for optimal memory. Sapolsky and colleagues (1986) showed that short-term acute stress leads to transient receptor loss, and speculate that prolonged chronic stress may produce permanent degeneration of hippocampal neurons, thereby affecting memory processes.

The hippocampus and pre frontal cortex (PFC) both play integral roles in memory and both have dense concentrations of glucocorticoid receptors. It is hypothesized that stress (and consequent increase in glucocorticoid levels) can impair contextual and episodic memory tasks that are known to require hippocampal and PFC function (Payne et al., 1996).Unsurprisingly, therefore, numerous studies have shown that hippocampus-based forms of memory are particularly affected by stress (e.g., Lupien et al., 1997).

IMPACT OF STRESS ON VARIOUS KINDS OF MEMORY

Although it is widely accepted that psychological stress can increase HPA axis activity, not all memory systems are equally affected by stress and cortisol levels. For instance, hippocampal-dependent forms of memory, particularly declarative memory, are affected by increased cortisol levels, whereas non-declarative forms of memory, such as procedural memory, appear to be unaffected. Furthermore, verbal declarative memory is impaired with increasing cortisol levels, whereas non-verbal declarative memory seems to be unaffected (e.g., Lupien et al., 1997, 1999).

Numerous experiments also indicate that working memory is negatively affected by increases in cortisol levels (Wolf, Schommer, Hellhammer, McEwen, & Kirschbaum, 2001) and that, in fact, this form of memory may be more sensitive to increases in cortisol levels than declarative memory (Lupien et al., 1999). For instance, Lupien and colleagues (1999) found that acute doses of corticosteroids caused significant decreases in working memory function, without significant changes in declarative memory. Finally, a decrease in performance on a variety of spatial memory tasks following exposure to increased cortisol levels has been observed in numerous experiments (Bonito Attwood, 2008; Luine, Villages, Martinex, & McEwen, 1994; Schwabe et al., 2007).

Stress and false memory

Only one study has examined the effect of stress on false memory. Payne and colleagues (2002) used the DRM paradigm to elicit false memories in stressed and non-stressed participants. Their results indicated that, within a DRM-type experimental paradigm, stressed individuals made significantly more false memory errors than did non-stressed controls, with the former finding it more difficult to distinguish between presented words and non-presented lures. Furthermore, they found that (a) while stress did increase false memory, it did not affect accuracy of memory for presented words, and (b) non-stressed participants responded significantly more quickly when correctly recognising presented items compared to when they incorrectly recognised critical lures; no such distinction was found in the stressed participants.

SUMMARY

It appears then that false memory recognition rates are affected by the kinds of materials that one is intended to remember, as well as to the presence of an environmental stressor. Furthermore, it is not clear whether errors of memory retrieval originate during encoding, consolidation, or retrieval processes. Although a large literature exists on the impact of stress on memory, and the impact of material specificity on false memory, until now no study has simultaneously looked at the impact of these two factors on false memory.

SPECIFIC AIMS AND HYPOTHESES

It is widely accepted that the hippocampus is both affected by cortisol and involved in episodic and declarative memory processes. This study aimed to determine whether acute stress impacted two different types of false memory: a) false memory for words, and b) false memory for pictures. Furthermore, it assessed the decay of both true and false memory over a 24-hour period. This is the first study to investigate the effects of stress on the material specificity of false memory.

More specifically, this study aimed to replicate experiment 1 from Gallo and colleagues 2004 study, adding stress and time retention as two new independent variables. Whereas Gallo et al.'s study only tested participant's memory immediately after they had studied the original word/picture lists, the current study featured memory tests of the original lists both immediately after and 24 hours later. This procedure enabled me to investigate the decay of true and false memory (for both pictures and words) over time. True and false memory differ in number of ways, one of which being the rate at which they decay over time. Payne et al. (1996) found that true memories in a recognition test decline with increasing time delay, whereas false memories remain relatively stable over a 24-hour delay. Similarly, other studies have found that true memory in recall tests are more affected by increasing retention intervals than are false memories (e.g., McDermott, 1996). Gallo's study (2004) did not look at the effect of retention time on the decay of memory, and this effect has never been studied with reference to the material specificity of false memory. Furthermore, Gallo's study did not look at the effect of stress on the material specificity of false memory.

The main hypotheses for this study were as follows:

1) False memory (for both pictures and words) in stressed participants will be greater than false memory (for both pictures and words) in non-stressed participants.
2) False memory for words will be greater than false memory for pictures in both stressed and non-stressed participants.
3) False memory rates will not increase over the 24-hour retention period in both stressed and non-stressed participants.
4) True memory rates will decrease over the 24-hour retention period, with stressed participants showing a larger decrease than non-stressed participants.

DESIGN AND METHODOLOGY

Design

This study was a true experimental, cross-sectional, 2 x 2 factorial design. It compared a specific cognitive process (false memory for both word lists and pictures) in two groups of subjects: one that had been exposed to an acute psychosocial stressor, and the other that had not. Additionally, each group was composed of an equal numbers of males and females. The independent variables in this study were stress manipulation (or lack thereof) and time (whether the memory tests were given immediately after or 24-hours later), with the two dependent variables being false memory for words and false memory for pictures.

Participants

Forty-three undergraduate students (21 male, 22 female) from the University of Cape Town's Department of Psychology were recruited for this study; they participated in exchange for course credit. Exclusion criteria included the presence of current psychoactive medication and current psychopathological conditions, and a history of neurological insult. All participants were between the ages of 18 and 35 years.

These inclusion/exclusion criteria are typical of studies into the effects of stress on cognition. With regard to age, elderly individuals show higher cortisol levels than younger indivduals (Kuldieka & Kirschbaum, 2005). Furthermore, numerous studies have shown that hippocampal neurons are lost with age (e.g., Bodnoff et al., 1995) and in individuals who have chronic illness in which elevated cortisol levels are present (e.g., Kahn, Rubinow, Davis, Kling, & Post, 1988). This could partly explain why elderly individuals and individuals with certain pathological conditions (such as depression or Cushing's disease) display poorer memory functioning (Kirschbaum et al., 1996).

With regard to the exact recruitment procedures, participants put their names on sign-up sheets posted in the Department of Psychology. Females were enrolled in the study if they were not taking any oral contraceptives and reported having a regular menstrual cycle (30 days). Females who remembered the precise dates of their previous menstrual cycle were given an appointment 6 days before the first day of their next menstrual cycle (to ensure they were in the late luteal phase of the menstrual cycle).[1]

Potential female participants who did not remember the exact dates of their menstrual cycle were asked to contact the experimenter on the first day of their next period. An appointment was then set up in a similar manner as described above. Menstrual cycle phase was checked post-experiment by participants emailing the experimenter.

Participants were pseudo-randomly assigned to either a Stress group or a Control group to ensure equal numbers of males and females in each group. For instance, if a pair of male and female participants was assigned to the Stress group, the next pair was assigned to the Control group. In summary, the initial sample size was 43 participants: Stress group $n = 22$ (11 males and 11 females), and Control group $n = 21$ (10 males and 11females).

Materials

Depression Screening Measure

The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) is a 21-item self-rated multiple-choice instrument that was developed to measure the intensity, severity, and depth of depression in patients. Higher ratings indicate greater symptom severity and more intense depression.

The BDI-II has been shown to be a reliable measure of depression in numerous studies and clinical settings (e.g., Beck, Steer, & Garbin, 1988). It possesses high internal consistency and correlates positively with other depression measures (Weeks & Heimberg, 2005).

Numerous studies have reported that patients suffering from major depressive disorder and depression in general, show elevated cortisol levels and have different phase shifts in adreno-cortical functions compared to people with no current mood disorder (Kirschbaum et al., 1996; Kudielka & Kirschbaum, 2005; Sapolsky et al., 1986). In the current study, the BDI-II was used to screen for depression, with severely depressed participants (BDI-II score of greater than 29) being excluded on the basis that their baseline cortisol levels differ from everyone else's.

Self-Reported Anxiety

The Spielberger State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) consists of two 20-item self-report scales, with each item having four possible answers. The STAI measures a person's state and trait anxiety separately. The 20-item State scale requires the respondent to describe the intensity of his/her feelings of anxiety at the current time. The 20-item Trait scale requires the respondent to describe the frequency with which they generally experience anxiety-related symptoms.

Psychometric studies indicate that the scale has a high degree of internal consistency (alpha = 0.92), as well as high test-retest reliability. In addition, the STAI correlates positively with the Taylor Manifest Anxiety Scale and the IPAT Anxiety Scale, both of which are reliable measures of anxiety levels (Spielberger & Vagg, 1984). The Trait scale was used to assess participants' general anxiety levels, while the State scale was used to assess participants' subjective experiences of anxiety throughout the experiment.

Physiological Measures

As in previous studies (e.g., Bonito Attwood, 2008; Schwabe et al., 2007), heart rate and saliva cortisol measurements were taken as objective measure of stress levels (e.g., Kirschbaum, Pirke, & Hellhammer, 1993). Only the latter are reported here.

Saliva samples were collected using Salimetrics Eyespear Sorbettes. This collection method has been successfully used in a previous study in our laboratory (Bonito Attwood, 2008).

For the current study, participants were instructed to place the cellulose-cotton eyespear under their tongues for 1 minute. After removal, the eyespear was placed into a conical tube, immediately stored in the laboratory's freezer, and eventually transported to an accredited laboratory for cortisol analyses. Assessment of cortisol in saliva has proven a valid and reliable reflection of the unbound hormone in the blood. Saliva cortisol measurements as an objective measure of stress have numerous advantages over blood cortisol measurements, including: stress-free sampling, lab independence, lower costs, non-invasive collection methods and the ability to obtain an unlimited frequency of measurements (Kirschbaum & Hellhammer, 1994). In addition, blood cortisol measurements are not always reliable measures of free cortisol levels (Kirschbaum et al., 1999).

The Acute Social Stressor: The Trier Social Stress Test (TSST)

The Trier Social Stress Test (TSST) is a highly standardized and widely used laboratory test used to induce psychosocial stress (for a full description of the procedure, see Kirschbaum et al., 1993). Compared with other laboratory-based stress induction tasks, the TSST provokes the most reliable and robust physiological stress (Dickerson & Kemeny, 2004). Six independent studies reported it producing a 2-4 fold elevation in salivary cortisol levels, with consistent increases in ACTH concentration and heart rate across different populations (Kirschbaum et al., 1993). Furthermore, a large number of studies have reported that laboratory tasks such as public speaking and mental arithmetic (both of which are included in the TSST) can increase cortisol levels (Het & Wolf, 2007; Kuhlmann, Piel, & Wolf, 2005). Finally, a meta-analysis reviewing conditions capable of evoking increased cortisol responses found that motivated performance tasks elicited the largest cortisol and ACTH responses if they were uncontrollable or characterized by social evaluative threats (Dickerson & Kemeny, 2004). The TSST is a motivated performance task, is uncontrollable, and contains an element of social evaluation, which would explain why it is so good at eliciting a stress response.

The TSST version used in this experiment was slightly modified from the original version described by Kirschbaum et al. (1993), but was identical to that used, with some success, by Bonito Attwood (2008) (see Bonito Attwood, 2008 for a full explanation of the changes made to the original TSST).

In accordance with the original TSST procedure, a participant was read a set of standard instructions, which are designed to introduce him/her to the task of the TSST. Participants were asked to assume the role of a job candidate for a job of their choice, and then given 10 minutes to prepare a speech detailing their suitability for that job. After the 10 minute preparation period, the participants in the stress group were given 5 minutes to deliver their speech. If the participant stopped speaking before time was up, the researcher said, "You still have time left, please continue." If the participant was still unable to continue delivering the speech, this set of standard questions was be asked: 1. "Please tell us what are some of your weaknesses"; 2. "What is the most difficult experience that you have had that would help you on the job?"; 3: "For what reasons should we not take you?"

After completion of the speech, participants were asked to perform a serial subtraction task, in which they started at the number 1022 and kept subtracting 13 until they were told to stop by the researcher. Each incorrect subtraction required the participant to start again at 1022. This mental arithmetic task lasted a full 5 minutes, as in the original TSST procedure.

The False Memory Task

The false memory test that was used in this study is an exact replication of the one used by Gallo and colleagues (2004) in Experiment 1 of their study. Study materials consisted of 288 unrelated common words (average word length was 6.1 letters), presented in black font on a white background (see Appendix B). Each of those words (e.g., *house*) was followed either by a picture (e.g., a picture of a house), or the same word printed slightly larger in a red font. Some of the black words were presented once (either followed by a corresponding picture or a red word), and others were presented twice (once followed by a picture, and once followed by a red word). To prevent sequencing effects, four counterbalancing conditions were created. This ensured that every fourth subject received a different study and recognition test. The study and recognition tests were equally distributed among the two groups (Stress and Control). Furthermore, the two criterial recollection tests were counterbalanced across participants, resulting in total of four counterbalancing conditions.

Study materials were presented via PowerPoint slides, and test materials via computer using E-Prime software (Version 1.1, Psychology Software Tools, Inc., Pittsburgh, PA, 2002). Participants studied 216 unique items, with 1/3 presented as red words, 1/3 as pictures, and 1/3 as both red words and pictures. Each studied item was first presented in black lowercase letters using Courier font for 700 ms. The black word was then replaced by either a picture, or by the same word in red-coloured Eros Bold ITC font (visibly larger than the Courier font) for 2000 ms. A 700-ms blank computer screen separated each picture or red word from the next study item. Items were randomly presented during the study phase, with those items presented as both pictures and red words randomly spaced throughout the study phase. The study phase took 16.5 minutes to complete.

After participants had learnt the study materials, they were given three recognition tests: a standard recognition test, followed by two criterial recollection tests. All words on the recognition tests were presented in the same black font used for the study items. Each recognition test contained items that had been studied (as either red words or pictures or, both) and nonstudied items. For the standard recognition test, participants were instructed to say “yes” to any item that had been studied (regardless of whether it had been presented as a red word, a picture, or both) and “no” to any items they felt were new/nonstudied. For this test, 3/4 of the items will be targets (original study material) and 1/4 lures (nonstudied items). The two criterial recollection tests were the red word test and the picture test. For the red word test, participants were required to say “yes” to any item they remembered studying as a red word. In addition, they were reminded that some red words were also studied as pictures (which they could still respond “yes” to). For the picture test, instructions were the same as for the red word test, except participants were instructed to say “yes” only if words had been studied as pictures (this could include words studied as both pictures and red words). For the criterial recollection tests, ½ the items were targets (present during the study phase) and ½ lures (not present in the study phase).

Procedure

Following conventions established by numerous studies (e.g., Kirschbaum et al., 1996), participants were tested between 12h00 and 16h00.[2]

The study procedures were completed over 2 days. On the first day all participants were treated exactly the same, regardless of group assignment. On the second day of testing, participants were treated differently depending on group assignment. The Stress group was subjected to the TSST procedure, whereas the Control group engaged in a 20-minute relaxation period.

A reminder phone call was made to participants one day before their Day 1 appointment. Upon arrival in the laboratory on Day 1, participants were given a consent form (see Appendix C), which gave them a brief outline of the study requirements and listed their rights as research participants. After reading and signing the consent form, participants were instructed to fill out the BDI-II and the STAI Trait. Following this, the false memory test was administered, which included the study phase and all three recognition tests. Participants were dismissed from the laboratory after a reminder to refrain from smoking, chewing gum, physical exercise, eating large meals, and drinking alcohol, fizzy drinks, tea or coffee 2 hours prior to their appointment on Day 2.[3]

On Day 2, participants were again tested between 12h00 and 16h00. Upon arrival, they were asked to complete a STAI State scale and a saliva sample was taken. Participants in the Stress group were then administered the TSST. Participants in the Control group were not administered any part of the TSST procedure. Instead, they relaxed in a room for 20 minutes, were seated in a comfortable chair and given non-political magazines (*Femina* and *Men's Health*) to read while relaxing music (Enya) played.

Following the 20-minute TSST and relaxation periods, participants in both groups were then instructed to relax for 5 minutes, after which a second saliva sample was taken. Participants were then instructed to complete the STAI State scale again. Following this, the false memory tests were administered, with participants completing the same 3 recognition tests they did on the previous day.

Finally, participants completed the STAI State scale, and a third saliva sample was taken. All participants were fully debriefed before they left, and the experimenter ensured that no participant in the Stress group was experiencing any distress due to the experimental procedures. Ethical approval had been granted for this study prior to data collection.

Data Analysis

Saliva samples were stored in a freezer within 30 minutes after collection. They remained there for the duration of data collection, after which they were delivered to National Health Laboratory Services at Groote Schuur Hospital for analyses.

The E-Prime software (Version 1.1, Psychology Software Tools, Inc., Pittsburgh, PA) used to present study materials to participants generated a unique data file for each participant after each recognition test. These data, as well as physiological measures and STAI scores, were the subject of subsequent statistical analysis.

Numerous participants had to be excluded from the final sample for various reasons. For instance, many participants delivered insufficient saliva samples for cortisol analysis, and some participants in the Stress and Control did not show manipulation-appropriate changes in cortisol levels after the experimental procedure. Additionally, some female participants began their period during the testing and had to be dropped. The final sample size was $n = 18$ (Stress group: $n = 6$ (3 female); Control group: $n = 12$ (6 female).

Initially, I investigated descriptive statistics, characterising the performance of the two groups on the Day 1 and Day 2 measures listed above. For the most part, between-group differences were assessed using individual samples *t*-tests. With regard to those significance tests, if Levene's test of equality of error variances was significant (indicating that the assumption of homogeneous variances across groups was not met), the *t*-test was run with a separate variance correction. The significance level for all tests was $\alpha = 0.05$.

RESULTS

Depression Screening

Scores for participants in the Stress (*M* = 8.5, *SD* = 7.45) and Control (*M* = 8.92, *SD* = 4.21) groups were not statistically significantly different on this measure, *t*(16) = -.15, *p* = .880 *d* = .08. The mean scores for each of the groups fell in the range conventionally described as "minimally depressed" (Beck et al., 1996). With regard to mood, then, it appears that the participants were representative of the general population and that neither group contained significantly more depressed participants than the other.

Measures of Stress

All analyses for the measures of stress were two-tailed, unless otherwise specified.

Trait Anxiety. Table 1 presents the participants' self-reported trait anxiety scores. Participants in the Stress group and Control groups were not significantly different on this measure, *t*(16) = -.09, *p* = .933 *d* = .04. With regard to anxiety levels, then, it appears that participants in both groups were equivocal as they entered the experiment.

To ensure that participants in the current sample were representative of the general population in terms of trait anxiety, I compared their scores to normative data for college students presented in the STAI test manual (Spielberger et al., 1983). Male participants (*n* = 9; *M* = 37.11, *SD* = 9.65) were not significantly different from the normative male population (*M* = 38.30, *SD* = 9.18), *t*(8) = - .37, *p* = .721. Female participants (*n* = 9; *M* = 38.44, *SD* = 0.93) were also not significantly different from the normative female population (*M* = 40.40, *SD* = 10.15), *t*(8) = - .60, *p* = .567. These results suggest that, with respect to trait anxiety, the current sample was representative of the general population of individuals of their similar age and education.

State Anxiety. With regard to self-reported state anxiety at the beginning of the experimental protocol (i.e., before the stress manipulation or relaxation period), participants in the Stress and Control groups were not statistically significantly different, *t*(16) = - .48, *p* = .639 *d* = .24. This result confirms that participants entered the experiment in the same general state of mind.

Participants in the Stress group reported an increase in state anxiety from pre-TSST to post-TSST, whereas participants in the Control group reported a decrease in state anxiety from pre-relaxation to post-relaxation (see Table 1). A set of two separate dependent samples t-tests confirmed that the change was significant for the Control group, $t(11) = 3.46$, $p = .003$, $d = .75$(one-tailed), but not for the Stress group, $t(5)= -1.35$, $p = .117$, $d = .31$ (one-tailed).

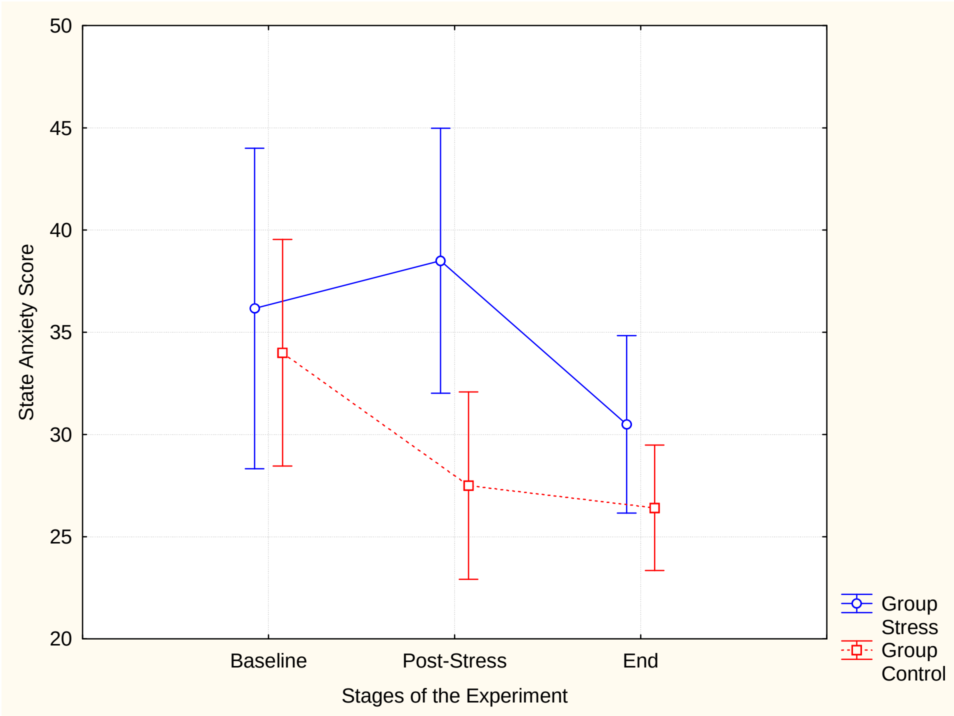

Figure 1. Self-reported state anxiety during the experiment.
Note. Vertical bars denote 0.95 confidence intervals

From an ethical standpoint, it was important to know whether the participants departed the laboratory in approximately the same state of mind as when they arrived. Self-reported levels of state anxiety did appear to differ from at the end of the experimental protocol compared to the start of the session; however, this difference was a decrease in anxiety (see Table 1). A set of two separate dependent samples t-tests, comparing state anxiety measured at baseline and at the conclusion of the experimental procedures (i.e., immediately before the participants were debriefed) confirmed that participants left the experiment with a lower level of anxiety than when they entered. For both the Stress and Control groups there was a statistically significant effect of the experimental procedures, $t(5)= 4.83$, $p = .005$, $d = 1.17$ and $t(11)= 3.33$, $p = .007$, $d = .92$, respectively.

Cortisol Levels

Regarding cortisol levels at the beginning of the experimental protocol, participants in the Stress and Control groups were not significantly different, $t(16) = -.62$, $p = .545$, $d = .31$. This result confirmed that participants entered the experiment with similar cortisol levels and HPA-axis activity.

In the Stress group, average free cortisol increased in response to the TSST from 2.67 ± 2.30 nmol/1 to 4.91 ± 4.15 nmol/1 (see Figure 2). A dependent samples t-test showed there was significant main effect of the TSST, $t(5) = -2.49$, $p = .027$, $d = .57$ (one-tailed). In the Control group, average free cortisol decreased in response to the relaxation period from 3.39 ± 4.09 nmol/1 to 1.66 ± 1.68 nmol/1 (see Figure 2). A dependent samples t-test showed there a significant main effect of the relaxation period, $t(11) = 2.40$, $p = .018$, $d = .55$ (one-tailed).

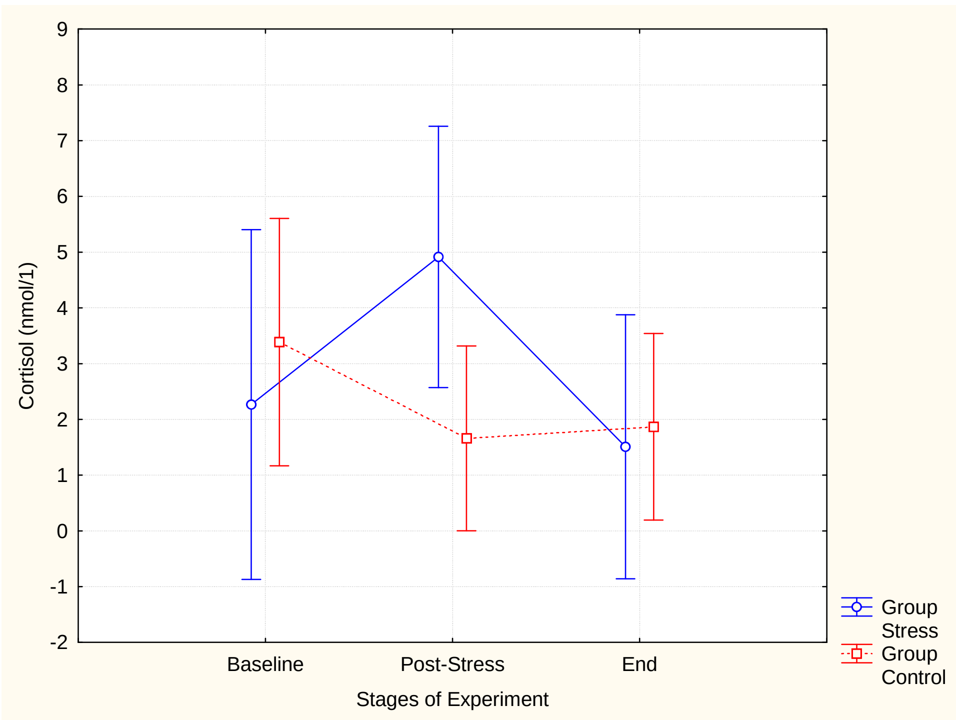

Figure 2. Salivary cortisol levels during the experiment.

Note. Vertical bars denote 0.95 confidence intervals.

The above results confirm that the TSST worked as expected to raise cortisol levels in the Stress group, and that the relaxation period was effective in decreasing cortisol levels in the Control group.

Recognition Tests

A summary of the Stress and Controls groups' performance on each recognition test is given in Table 2. All subsequent analyses will be one-tailed (unless otherwise specified) as most of the statistical tests relied on directional hypotheses.

Day 1 Analyses

For the initial analyses, participants were not split into groups as all were treated identically on Day 1 (See Table 3 for combined scores). Furthermore, the analysis of Day 1's results were intended to see whether they were consistent with Gallo's (2004) findings, and for the most part individual *t*-tests were done so as to replicate the analyses conducted in that paper.

The first analysis investigated whether items studied as both pictures and red words were recognized better than items presented as either pictures or words, and if (following picture superiority effect predictions derived from distinctiveness heuristic theories) items studied as pictures only were recognized better than items studied as red words only. A one-way ANOVA on data from the Standard recognition test revealed a significant main effect of Item Type, $F(2,51) = 15.10$, $p < .001$. Post-hoc tests (Tukeys HSD) showed that the number of hits (correct responses) for items studied as both pictures and red words (i.e., Both hits) was significantly greater than the number of hits for items studied only as pictures (i.e., Picture hits) or for items presented only as red words (i.e., Red Word hits), $p = .002$ and $p < .001$, respectively (See Figure 3). The data showed a picture superiority effect, in which the number of Picture hits achieved by participants was statistically significantly greater than the number of Red Word hits, $p = .049$. This result is consistent with the distinctiveness heuristic, whereby pictures are more likely to be remembered than words due to their more distinctive perceptual qualities.

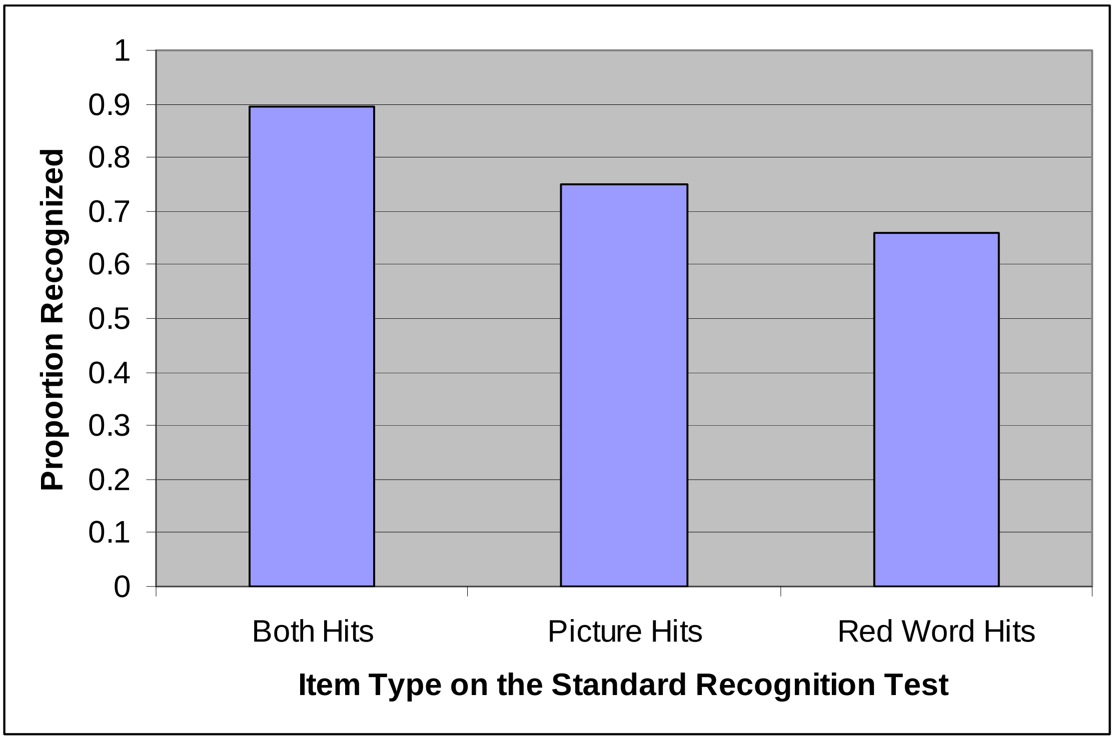

Figure 3. Comparison of correct responses for Day 1 on the standard recognition test

To further investigate whether the picture superiority effect was demonstrated by the current data, I compared the number of Picture hits on the Picture test against the number of Red Word hits on the Red Word test (See Table 4). The number of hits for items presented as pictures on the Picture test was not statistically significantly greater than the number of hits for items presented as Red Words on the Word test. The distinctiveness heuristic prediction here is that the former would be greater than the latter, due to the more distinct perceptual qualities of pictures compared to words.

To confirm that items studied twice (i.e., as both pictures and as red words) were recognized better throughout the Day 1 recognition tests, separate independent-samples *t*-tests were conducted on data from the criterial recollection tests. The results of those analyses are shown in Table 4. As can be seen, no statistically significant results were found when comparing participant's hits for items presented as both pictures to words to hits for items presented as pictures only on the Picture test. One explanation for this non-significant finding is that, although items presented as both pictures and words were likely to be very familiar, items presented as pictures only should also have been familiar (following the distinctiveness heuristic). On the Red

Word test, however, hits for words were significantly greater than hits for items presented as both pictures and words. This is a curious result, as items presented as both pictures and words should have been more familiar as they were presented twice, and should therefore have been more likely to be remembered.

To investigate false memory on Day 1, separate sets of independent-samples *t*-tests were run to compare the average number of false alarms (FAs) in each of the two criterial recollection tests (see Table 5). On the Red Word criterial recollection tests, Picture FAs were not significantly greater than New FAs. Similarly on the Picture criterial recollection test, Red Word FAs were not significantly greater than New FAs. These results do not match the predictions I made on the basis of Gallo et al.'s (2004) work: New FAs are false alarms for items that were not presented to participants during the study phase; they should therefore have been less familiar to the participants and thus less likely to be (falsely) remembered.

Consistent with a prediction deriving from the picture superiority effect, however, New FAs were significantly lower on the Picture test than on the Red Word test (see Table 5). Moreover Red Word FAs on the Picture test were significantly lower than Picture FAs on the Red Word test. The lower amount of FAs on the Picture test suggests that studying and recalling pictorial materials, which are more distinctive than words, leads to fewer false memory errors.

Day 2

To further investigate the picture superiority effect, separate sets of independent-samples t-tests were run to compare Picture and Word hits within the Standard recognition test, and between the two Criterial recollection tests in both the Stress and Control groups. Results shown in Table 6 indicate that stressed participants made significantly more Picture hits compared to Red Word hits on the Standard test, as did the Control group (See Table 7). This results supports the picture superiority effect, whereby pictures are more likely to be remembered than words (due to predictions made by the distinctiveness heuristic).Both the Stress and Control group (See Table 6 and 7 respectively) did not make significantly more Picture hits compared to Red Word hits between the two Criterial recollection tests (i.e., Picture hits on Picture test were not significantly greater than Red Word hits on the Red Word test). This result is does not support the picture superiority effect, and cannot be explained by the distinctiveness heuristic as picture should always be better remembered than words.

Hypothesis One stated that participants in the Stress group would commit more false memory errors on all of the recollection tests than would participants in the Control group. The Day 2 analyses therefore focused on between-group comparisons of performance on the three recognition tests (specifically, independent samples *t*-tests, with number of FAs as the dependent variable) to detect whether participants performed statistically significantly differently after their respective experimental manipulations. As shown in Table 8, New FAs on all three recognition tests did not differ significantly between the Stress and Control group. As shown in Figure 4, it can been seen that while differences between the two groups were not significant, the Stress group tended to show greater false memory scores on both the Picture and Red Word criterial recollection tests. This is in support of the hypothesis that false memory would be greater in the Stress group.

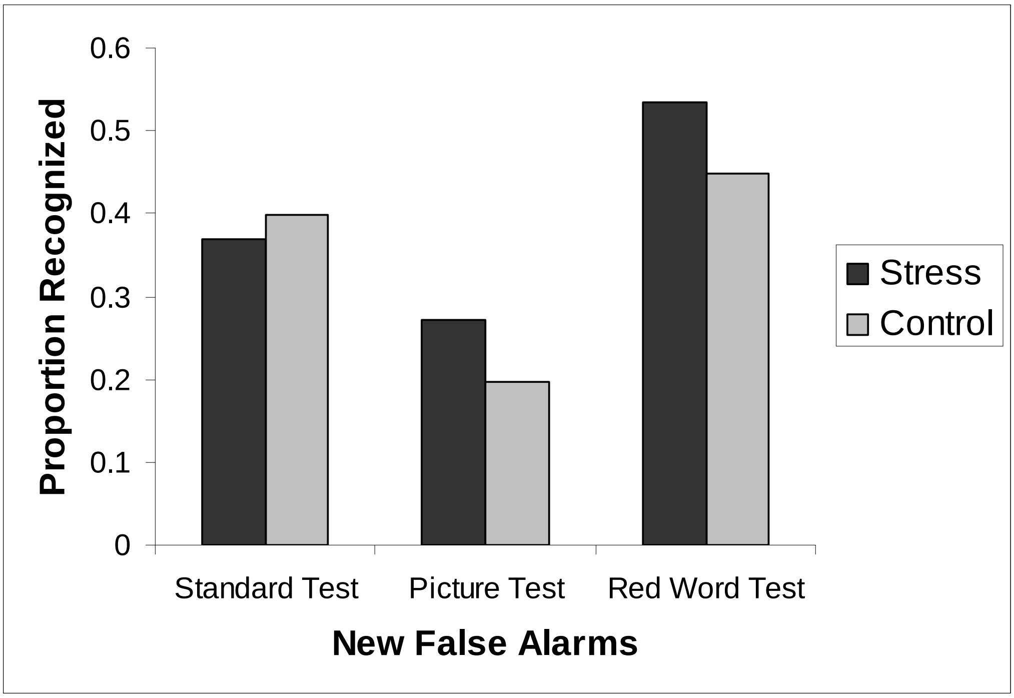

Figure 4. Comparison of New False Alarms Scores between the Two Groups on each Recognition Test

To further investigate whether stressed participants displayed more false memory errors, a set of two separate independent groups *t*-tests were run comparing Picture FAs and Red Word FAs on the two criterial recollection tests. Results in Table 8 indicated that the number of Red Word FAs made by participants on the Picture test did not differ significantly between the two groups. A similar result was found for Picture FAs on the Red Word tests, with the two groups not showing significantly different scores.

In addition to the hypothesis that the Stress group would commit more false memory errors, hypothesis Two predicted that false memory for words would be greater than false memory for pictures in both the Stress and Control group. To investigate this, separate independent-samples t-tests were run on each group to compare Picture FAs on the Red Word Test to Red Word FAs on the Picture Test. Results shown in Table 9 indicate that non-stressed participants made significantly more Picture FAs than Red Word FAs. This result is not consistent with the hypothesis that false memory would be higher for words compared to pictures. Moreover, results indicated that there was no significant difference between the amount of Red Word FAs and Picture FAs within the stressed participants. However, this non-significant result tended towards significance in the desired direction (as participants made more Red Word FAs than Picture FAs) and could merely be due to the small sample size. The latter result supports the hypothesis that false memory for words would be greater than false memory for pictures.

Comparison of Day 1 and Day 2 scores

Hypothesis Three predicted that false memory rates would remain stable over a 24-hour retention period. To investigate this, sets of separate dependent-samples t-tests were used to compare the within-groups performance on the three different recognition tests with regard to false memories (false alarms).

Within the Stress group, New FAs did not differ significantly between Day 1 and 2 on the Standard test, Red Word test, or Picture test (see Table 10). Furthermore, Picture FAs on the Red Word test did not differ significantly between Day 1 and 2, nor did Red Word FAs on the Picture test. Within the Control group, New FAs did not differ significantly between Day 1 and 2 on the Standard test, Red Word test, or Picture test (see Table 11). Furthermore, Red Word FAs on the Picture test did not differ significantly between Day 1 and 2, however, Picture FAs on the Red

Word test did differ significantly. Most of the above results support the hypothesis that false memory error rates would remain stable over a 24-hour retention period.

Hypothesis Four predicted that true memory rates would remain decrease over a 24-hour retention period. To investigate this, sets of separate dependent-samples t-tests were used to compare the within-groups performance on the three different recognition tests with regard to true memories (i.e., hits).

Within the Stress group, hits for items presented as both pictures and words did not differ significantly between Day 1 and 2 on the Standard test, Red Word test, or Picture test (see Table 10). However, results on the Picture test were almost significant and in the desired direction (showing a decrease from Day 1 to Day 2), therefore a larger sample size would probably yield more powerful results. Furthermore, Picture hits on the Standard test and Picture test did not differ significantly between Day 1 and 2. Again, results tended towards significance in the desired direction, in support of the hypothesis that true memory decreases over a 24-hour retention period. Lastly, Red Word hits on the Standard test and Red Word test did not differ significantly between Day 1 and 2. These results do not confirm Hypothesis Four, yet some results tend towards significance, indicating that a larger sample size is needed.

Within the Control group, hits for items presented as both pictures and words did not differ significantly between Day 1 and 2 on the Red Word test and Picture test, but did differ significantly on the Standard test (see Table 11). The latter result is in support of the hypothesis that true memory decreases over a 24-hour retention period. However, Picture hits on the Standard test and Picture test did not differ significantly between the two days. Furthermore, Red Word hits on the Standard test and Red Word test did not differ significantly between the two days. These results are not consistent with the hypothesis that true memory would decrease over a 24-hour retention period.

Difference Scores (Day 1 – Day 2)

Difference scores were calculated for both the Stress and Control group by taking their Day 1 scores minus their Day 2 scores (Table 12).

Hypothesis 4 predicted that true memory would decrease over a 24-hour retention period, but furthermore, that this decrease would be greater in the stressed participants. To investigate this, separate sets of independent samples *t*-test were run to directly compare the Stress and Control group's difference scores on each recognition test for true memory (hits).

On the Standard recognition test, difference scores for hits on all items (Both hits, Picture hits, and Red Word hits) did not differ significantly between the Stress and Control group (see Table 13). These results do not confirm the hypothesis that the Stress group would show a more significant decrease in true memories over a 24-hour retention period.

Results in Table 13 indicate that when comparing difference scores on the Red Word test, participants hits for items presented as both pictures and words during the study phase and hits for items presented as Red Words only did not differ significantly between the Stress and Control group. Again, these results contradict the hypothesis that the Stress group would show a more significant decrease in true memory over a 24-hour retention period.

Results in Table 13 also indicated that when comparing difference scores on the Picture test, participants hits for items presented as both pictures and words during the study phase, and hits for items presented as Pictures only did not differ significantly between the Stress and Control group. Again, these results contradict the hypothesis that the Stress group would show a more significant decrease in true memory over a 24-hour retention period.

DISCUSSION

Based on evidence from numerous studies, this study predicted that acute psychosocial stress would temporarily cause disrupted hippocampal and PFC functioning in the Stress group, thereby increasing the likelihood of false memories. Results pertaining to a check of the experimental manipulation indicated that the administration of the TSST was successful in significantly raising cortisol levels and self-reported anxiety in Stress group participants. Furthermore, participants in the Control group showed significantly lowered cortisol levels and self-reported levels of anxiety, following a period of relaxation. Therefore, as the participants entered the cognitive testing phase of the experiment, participants in the Control group were in a different physiological state to those in the Stress group, with the latter more likely to have temporarily impaired hippocampal function.

As mentioned in the results section, the analyses of day 1's results were to confirm that this study replicated the results found by Gallo (2004). On the Standard test a picture-superiority effect was found, as Picture hits were significantly greater than Red Word hits. This result can be explained by the distinctiveness heuristic, as pictures should always be better remembered than words due to their more distinctive perceptual qualities. A familiarity model could also explain this effect, as pictures should be more familiar that words due to their distinctive features, therefore, the more familiar pictures should be better remembered than words. When further investigating the picture superiority effect by directly comparing the two criterial recollection tests, it was found that Picture hits on the Picture test were not significantly greater than Red Word hits on the Red Word test. This result was not found by Gallo, and cannot be explained by the familiarity model or distinctiveness heuristic. Pictorial material should always be better remembered than word (due to the distinctiveness heuristic).

Comparisons were also run to investigate whether items presented as both pictures and words (Both hits) during the study phase were more likely to be remembered than items presented as Pictures or Red Words only. On the Standard recognition test it was found that Both hits were significantly greater than Picture hits and Red Word hits. Gallo found that on the Standard recognition test, Both hits were significantly greater than Red Word hits but not significantly greater than Picture hits. Items presented twice (as both pictures and words during the study phase) were more familiar to participants and should therefore be better remembered than items presented as Red Words only. While items presented twice during the study phase are familiar, pictures are also familiar as they are more distinctive (as explained by the distinctiveness heuristic). Therefore there should be no significant difference between hits for items presented as both pictures and words and items presented as only pictures (seen are they are both familiar, and therefore more likely to be equivocally remembered). Therefore it was expected that there would be no significant difference between Both hit and Picture hits. The current study did not find this on the Standard recognition test, but found it on criterial recollection tests. Both hits were significantly greater than Red Word hits on the Red Word test, but not significantly greater than Picture hits on the Picture test. This result is also consistent with Gallo’s findings.

Investigations were also run on false memory scores for Day 1. This study found that there was no significant different between the Picture FAs and New FAs on the Red Word test. This result cannot be explained by the familiarity model, as Picture FAs should still be more familiar than new items that were never been presented during study phase. Participants should be affected by the prior presentation of the to-be-excluded items (Picture FAs) and therefore more likely to remember them. The current study also found that Red Word FAs and New FAs on the Picture test did not differ significantly. Both of the above are not consistent with Gallo's results.

Most importantly, all false alarms were lower on the Picture test than on the Red Word test (as found by Gallo). Red Word FAs on the Picture test were significantly lower than Picture FAs on the Word test. Similarly, New FAs on the Picture test were significantly lower than New FAs on the Word test. These recollections are consistent with the distinctiveness heuristic as participants should expect more distinct recollections on the Picture test, thereby lowering all false alarms. A familiarity-based model could also explain these results. Seen as pictures were 'stronger' in memory than words, we could predict a more conservative criterion recollection response on the Picture test, thereby lowering the number of false alarms relative to the Red Word test.

To further investigate the picture superiority effect, separate comparisons were run on Day 2 results for the Stress and Control group. Results indicated that on the Standard recognition test, hits for items presented as pictures were significantly greater than hits for items presented as words, in both groups. This result demonstrates the picture superiority effect, whereby pictures are more likely to be remembered than words. However, hits for Pictures in the Picture tests and Red Word hits on the Red Word test were not significantly different in both the Stress and Control group. The latter result is not consistent with the distinctiveness heuristic, as pictures should always be better remembered than words.

Based on the distinctiveness heuristic (Gallo, 2004), it was predicted that false memory for words would be greater than false memory for pictures in both the Stress and Control group. When comparing picture FAs on the Red Word test to Red Word FAs on the Picture test, no significant difference was found between the two in the Stress group. Moreover, picture FAs were higher than word FAs in the Stress group. This could be explained by the fact that participants used more conservative response criterion in the Picture test, thereby lowering the word FAs. In the Control group, picture FAs were significantly greater than word FAs. This shows that false memories were higher on the word test compared to the picture test. However, the distinctiveness heuristic cannot explain these results as pictures should always be better remembered than words.

The current study hypothesized that false memories would be greater in the Stress group compared to the Control group. The between-groups analyses of Day 2's results did not indicate a significant difference on any of the recognition tests; however the Stress group did tend to make more false alarms than the Control group. Due to the disruption of hippocampal functioning by the increased release of cortisol in the stressed participant's one should expect them to make more memory errors than the control participants, however the results do not indicate this. A possible reason for the non significant result is the small sample size, especially in the Stress group. The results are promising, and tend towards supporting the hypothesis. Therefore further studies with a larger sample size should yield more promising results.

This studies third hypothesis predicted that false memory would remain stable over a 24-hour retention period. When comparing New FA scores on the Standard test, there was no significant difference between Day 1 and 2 in the Stress group, however there was a significant increase in FAs from Day 1 to 2 in the Control group. The Stress group's results support the hypothesis that false memories remain stable over a 24-hour retention period; whereas the Control groups results contradict it. When comparing New FA scores on the Red Word test, there was no significant difference between Day 1 and 2 scores in both the Stress and Control group. The same results were found on the Picture test, supporting the hypothesis that false memories are relatively stable over a 24-hour retention period. Furthermore, when comparing Picture FAs on the Red Word test and Red Word FAs on the Picture test, there was no significant difference between Day 1 and 2's scores in both the Stress and Control group, further supporting the third hypothesis.

The final hypothesis predicted that true memory would decrease over a 24-hour retention period, and that this decrease would be greater in the stressed participants. When comparing hits (true memories) on all three recognition tests, not significant differences were found between Day 1 and 2's scores in both the Stress and Control group. This contradicts the hypothesis that true memories would decrease over a 24-hour retention period. Although results were not significant, true memory almost always decreased from Day 1 to day 2 (as can be seen in Table 2). These results are promising as they are in the predicted direction and a larger sample size may yield more powerful results.

Furthermore, when directly comparing the difference scores between the Stress and Control group, no significant differences were found. This result does not confirm the hypothesis that the Stress group would show a greater decrease in true memories. While the results were not statistically significant, in general the Stress group showed a greater decrease in true memories compared to the Control group and a larger sample size may allow more powerful comparisons to be made.

Limitations and Directions for Further Research

The current study set out to show that there may be contrasting effects of acute psychosocial stress on the material specificity of false memory. Although not all hypotheses were confirmed, some results tended toward statistical significance in the predicted direction, which indicates that there is continued promise in the study of the impairing effects of stress on cognitive function.

Several limitations of the current study need to be addressed by future researchers who wish to clearly outline the relationship between stress and false memory. Firstly, the current sample size was very small and the effects being studied require a larger group of participants. Some results tended towards significance in the desired direction, however failed to reach significance. A larger sample size and more diverse sample should yield promising results. In addition, given the fragility of salivary cortisol samples (as shown by the number of participants in the current study who had to be dropped due to insufficient cortisol analyses), collecting larger numbers of participants is imperative.

The current study relied on self-report of menstrual cycle phase. Although it was for the most part accurate, several female participants had to be dropped from the study as they began their menstrual cycle during the testing days. Future studies might add physiological measures of menstrual cycle to ensure improved accuracy. Additionally, future research might investigate the effects of time of menstrual cycle on cognitive performance following stress.

Finally, both basal cortisol levels and cortisol increases in response to the TSST were lower than those reported in other studies (e.g., Kirschbaum, 1993). The difference in magnitude of the increase can be explained by differences in the administration of the TSST, and future studies should adhere more strictly to the original TSST procedure. The differing basal levels cannot be easily explained as participants in the current study were in the same age range, and tested at the same time as those in previous studies. Future research should possibly employ heart rate and skin conductance measurements to gain a better representation of one's physiological and cognitive responses to stress.

Future studies regarding false memory and stress could possibly look at the effect of neutral and emotional materials in eliciting false memories, as studies have shown the false memory tends to increase for neutral stimuli but decrease for emotionally arousing materials (Payne, Jackson, Ryan, Hoscheidt, Jacobs, & Nadel, 2006). In addtion, the effect of gender interactions on false memory and stress could also be looked at. This question then remains open for future research.

REFERENCES

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation

Beck, A. T., Steer, R. A., & Garbin, G. M. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77-100.

Bodnoff, S. R., Humphreys, A. G., Lehman, J. C., Diamond, D. M., Rose, G. M., & Meaney, M. J. (1995). Enduring effects of chronic corticosterone treatment on spatial memory learning, synaptic plasticity and hippocampal neuropathology in young and mid-aged rats. *Journal of Neuroscience*, *15*, 61-69.

Bonito Attwood, C. (2008). *The impact of acute psychological stress on spatial cognition.* Unpublished master's thesis, University of Cape Town, Cape Town, South Africa.

Boyer, P., Phillips, J. L., Rousseau, F. L., & Ilivitsky, S. (2007). Hippocampal abnormalities and memory deficits: New evidence of a strong pathophysiological link in schizophrenia. *Brain Research Reviews, 54*, 92-112.

Budson, A. E., Dodson C. S., Vatner, J. M., Daffner, K. R., Black, P.M., & Schacter, D. L. (2005). Metacognition and false recognition in patients with frontal lobe lesions: The distinctiveness heuristic. *Neuropsychologia*, *43*, 860-871.

Budson, A. E., Droller, D. B. J., Dodson C. S., Schacter, D. L., Rugg, M. D., Holcomb, P. J., et. al. (2005). Electrophysiological dissociation of picture versus word encoding: The distinctiveness heuristic as a retrieval orientation. *Journal of Cognitive Neuroscience*, *17*, 1181-1193.

Budson, A. E., Sitarski, J., Daffner, K. R., & Schacter, D. L. (2002). False recognition of pictures versus words in Alzheimer's disease: The distinctiveness heuristic. *Neuropsychology*, *16*, 163-173.

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, *15*, 191-205.

Dab, S., Claes, T., Morais, J., & Shallice, T. (1999). Confabulation with a selective descriptor process impairment. *Cognitive Neuropsychology*, *16*, 215-242.

de Kloet, E. R., Oitzl, M. S., & Joëls, M. (1993). Functional implications of brain corticosteroid receptor diversity. *Cellular & Molecular Neurobiology*, *13*, 433-455.

de Kloet, E. R., Oitzl, M. S., & Joëls, M. (1999). Cognition: Are corticosteroids good or bad guys. *Trends in Neurosciences*, *10*, 422-426.

deQuervain, D. J. F., Roozendaal, B., Nitsch, R. M., McCaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of declarative memory in humans. *Nature Neuroscience, 3*, 313-314.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130,* 355-391.

Gallo, D. A., Kensinger, E. A., & Schacter, D. L. (2006). Prefrontal activity and diagnostic monitoring of memory retrieval: fMRI of the critical recollection task. *Journal of Cognitive Neuroscience*, 18, 135-148.

Gallo, D. A., & Roediger, H. L. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory & Language*, *47*, 469-497.

Gallo, D. A., Weiss, J.A., & Schacter, D. L. (2004). Reducing false recognition with critical recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory & Language, 51*, 473-493.

Grady, C. L., McIntosh, A. R., Rajah, M. N., & Craik, F. I. M. (1998). Neural correlates of the episodic encoding of pictures and words. *Proceedings of the National Academy of Sciences, 95*, 2703-2708.

Het, S., & Wolf, O. T. (2007). Mood changes in response to psychosocial stress in healthy young women: effects of pretreatment with cortisol. *Behavioral Neuroscience, 121*, 11-20.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.

Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Sciences*, *2*, 137-145.

Kahn, J. P., Rubinow, D. R., Davis, C. L., Kling, M., & Post, R. M. (1988). Salivary cortisol: a practical method for evaluation of adrenal function. *Biological Psychiatry*, *23*, 335-349.

Kirschbaum, C., & Hellhammer, D. H. (1992). Methodological aspects of salivary cortisol measurement. In C. Kirschbaum, G. F. Reeds, & D. H. Hellhammer (Eds.), *Assessment of hormones and drugs in saliva in biobehavioural research* (pp. 19-32). Seattle WA: Hogrefe & Huber.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology, 19,* 313-333.

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine, 61*, 154-162.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The Trier Social Stress Test - A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*, 76-81

Kirschbaum, C., Wolf, O. T., May, M., Wippich, W., & Hellhammer, D. H. (1996). Stress and treatment-induced elevation of cortisol levels associated with impaired declarative memory in healthy adults. *Life Science, 58*, 1475-1483.

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis response to stress: A review. *Biological Psychology, 69,* 113-132.

Kudielka, B. M., Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology, 29,* 983-992.

Kuhlmann, S., Piel, M., & Wolf, O.T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Jounal of Neuroscience*, *25*, 2977-2982.

Luine, V., Villages, M., Martinex, C., & McEwen, B. S. (1994). Repeated stress causes reversible impairments of spatial memory performance. *Brain Research, 639*, 167-170.

Lupien, S. J., Gaudreau, S., Tchiteya, B. M., Maheu, F., Sharma, S., Nair, N.P.V., et al. (1997). Stress-induced declarative memory impairment in healthy elderly subjects: Relationship to cortisol reactivity. *Journal of Clinical Endocrinology and Metabolism, 82*, 2070-2075.

Lupien, S. J., Gillin, C.J., & Hauger, R.L. (1999) Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: A dose-response study in humans. *Behavioural Neuroscience*, *113*, 420-430.

Marsh, E. J., McDermott, K. B., & Roediger, H. L. (2004). Does test-induced priming play a role in the creation of false memories? *Memory*, *12*, 44-55.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language, 35*, 212-230.

McDermott, K.B., & Watson, J.M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language*, 45, 160-176.

Mendl, M. (1999). Performing under pressure: Stress and cognitive function. *Applied Animal Behaviour Science*, *65*, 221-244

Multhaup, K. S., & Conner, C. A. (2002). The effects of considering nonlist sources on the Deese-Roediger-McDermott memory illusion. *Journal of Memory & Language, 47*, 214-228.

Newcomer, J. W., Selke, G., Melson, A. K., Hershey, T., Craft, S., Richards, K., et al., (1999). Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Archives of General Psychiatry, 56,* 527-533.

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing and recollecting events that never occurred. *Journal of Memory & Language*, 35, 261-285.

Payne, J. D., Jackson, E. D., Ryan, L., Hoscheidt, S., Jacobs, W. J., & Nadel, L. (2006). The impact of stress on neutral nad emotional aspects of episodic memory. *Memory, 14*, 1-16.

Payne, J. D., & Nadel, L. (2004). Sleep, dreams, and memory consolidation: The role of the stress hormone cortisol. *Learning and Memory*, *11*, 671-678.

Payne, J. D., Nadel, L., Allen, J. J. B., Thomas, K. G. F., & Jacobs, W. J. (2002). The effects of experimentally induced stress on false recognition. *Memory*, *10*, 1-6.

Psychology Software Tools, Inc. (2002). *E-Prime getting started guide (version 1.1)*. Pittsburgh, PA: Author.

Rajaram, S., & Roediger, H.L. (1993). Direct comparison of four implicit memory tests. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *19*, 765-776.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory & Language*, *35*, 76-100.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented on lists. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*, 803-814.

Sapolsky, R. M., Krey, L. C., & McEwen, B. S. (1986). The neuroendocrinology of stress and aging: The glucocorticoid cascade hypothesis. *Endocrinology Reviews, 7,* 284-301.

Schacter, D. L., Israel, L., & Racine, C. (1999). Supressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language*, *40*, 1-27.

Schacter, D. L., Koutstaal, W., Johnson, M. K., Gross, M. S., & Angell, K. E. (1997). False recollection induced by photographs: A comparison of older and younger adults. *Psychology & Aging*, *12*, 203-215.

Schwabe, L., Oitzl, M. S., Philippsen, C., Richter, S., Bohringer, A., Wippich, W., et al. (2007). Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learning & Memory, 14*, 109-116.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Memory and Learning, 6*, 174-215.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State–Trait Anxiety Inventory*. Palo Alto, TX: Consulting Psychologists Press.

Spielberger, C. D., & Vagg, P. (1984). Psychometric properties of the STAI: a reply to Ramanaiah, Franzen, and Schill. *Journal of Personality Assessment, 48*, 95-97.

Weeks, J. W., & Heimberg, R. G. (2005). Evaluation of the psychometric properties of the Beck Depression Inventory in a non-elderly adult sample of patients with generalized anxiety disorder. *Depression & Anxiety*, *22*, 41-44.

Wolf, O. T., Schommer, N. C., Hellhammer, D. H., McEwen, B. S., & Kirschbaum,C. (2001). The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology, 26*, 711-720.

Wolkowitz, O. M., Reus, V. I., Weingartner, H., Thompson, K., Breier, A., Doron, A., et al. (1990). Cognitive effects of corticosteroids. *American Journal of Psychiatry*, *147*, 1297-1303.

APPENDIX A
DRM Explanation

The Deese-Roediger-McDermott (DRM) paradigm provides a means of creating false memories and has successfully been used in numerous studies (e.g., Marsh, McDermott & Roediger, 2004). Subjects are tested on a list of semantically associated words, all related to a non-presented critical lure. When later tested, subjects recall and recognise the non-presented critical word with unusually high probabilities, often as high or greater than studied items (Gallo & Roediger, 2002; Roediger & McDermott, 1995). An example of a word list given is the DRM is as follows: bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, and drowsy (Roediger, 1995). All these words are semantically related to the critical lure (which is sleep). Although the word sleep (the critical lure) was never presented during the study phase, it was presented in the recognition test. Subjects recall the critical lure with a probability comparable to recall of items presented in the middle of a list, thought to represent recall from long term memory (McDermott, 1996). Numerous experiments replicating the DRM paradigm have reported similar results (e.g., Multhaup & Conner, 2002; Payne et al., 1996).

APPENDIX B

Words used in the false memory test

recordplayer	butterfly	suitcase	heart
mortarboard	saddle	plant	yoyo
toaster	rainbow	goat	duck
powerstrip	hockeystick	mug	cow
pig	sneakers	flamingo	dresser
spatula	penguin	tissues	nest
teddybear	skateboard	wheelchair	horse
cigar	television	apron	thumbtack
fishingrod	carrot	joystick	dollars
pelican	pear	table	corn
bandana	mushroom	cookie	wrench
wolf	snake	candle	clarinet
clock	necklace	overalls	bull
football	rabbit	beetle	snail
speakers	slide	pliers	screwdriver
backpack	sandals	battery	golfbag
saturn	pot	rooster	america
maracas	bulldozer	accordion	bow
octopus	peas	clothespin	banana
files	brain	staplegun	dna
scale	turkey	elephant	pogostick
hat	bandaid	cloud	tank
canon	refrigerator	strawberry	tree
house	tent	flower	guitar
zebra	hourglass	lizard	donut
hanger	sandwich	ladder	watch
icecream	microscope	vacuum	doorknob
whistle	deer	calculator	giraffe
stapler	dartboard	computer	roller
bat	clipboard	tuba	tie
kettle	dumptruck	glasses	socks
bench	vase	meter	trophy
whisk	net	racket	package
hammer	bottle	mirror	pretzel
lantern	bucket	candycane	turtle
	jackolantern	matches	caterpillar
	stroller		harp

tape	telephone	blimp	car
headphones	peanuts	owl	fish
bicycle	doll	bread	binoculars
hotairballoon	grill	crab	broom
pen	toilet	pipe	corkscrew
mouse	gavel	tiger	briefcase
knife	blinds	camel	spider
fan	shorts	drill	sofa
cat	pepper	screen	notebook
paintbrush	rollerskate	rhinoceros	cherries
airpump	hook	scrubbrush	pan
walrus	iron	pants	racecar
buggy	dragon	camcorder	alligator
leaf	windmill	scissors	register
hydrant	towel	microwave	safetypin
spraybottle	jacket	jar	trombone
sewingmachine	lightbulb	cake	medal
shelves	axe	carousel	rocket
footprints	pitchfork	stethoscope	wagon
train	shoppingbag	fireplace	bell
pumpkin	ring	violin	shovel
comb	gear	rope	crown
coconut	lemon	skis	sweater
lifevest	hairdryer	helmet	extinguisher
dinosaur	compass	acorns	lion
scoop	lighthouse	lamp	marble
thermos	fork	palette	skull
tire	peacock	handtruck	mixer
tomato	ovenmitts	babycarriage	chest
belt	flag	frog	flippers
lighter	buffalo	bus	camera
pancakes	ostrich	scarf	lobster
kite	airplane	pie	desk
pillow	bathtub	iceskate	orange
astronaut	helicopter	monkey	clip
satellitedish	pineapple	pacifier	sled
moviereel	brazier	watermelon	fryingpan
barrel	dice	mask	tincan
seahorse	compactdisc	balloons	panda
armadillo	policecar	lock	rocker
drum	bed	mailbox	shirt
trashcan	cactus	projector	handcuffs
grapes	squirrel	needle	seal
fox	sharpener	photocopier	microphone
book	hamburger	gumballs	rake
moose	horn	sink	worm
roledex	cheese	magnifier	plate
globe	submarine	wastebasket	flashlight

nail	boat	piano	

umbrella satellite razor colander eggs apple crutches ballhoop shuttlecock giftbox highchair chesspiece gun seashell dumbbell pencil scooter parrot cassette hotdog			

APPENDIX C

Consent Form

Informed Consent to Participate in Research and Authorization for Collection, Use, and Disclosure of Protected Health Information

This form provides you with information about the study and seeks your authorization for the collection, use and disclosure of your protected health information necessary for the study. The Principal Investigator (the person in charge of this research) or a representative of the Principal Investigator will also describe this study to you and answer all of your questions. Your participation is entirely voluntary. Before you decide whether or not to take part, read the information below and ask questions about anything you do not understand. By participating in this study you will not be penalized or lose any benefits to which you would otherwise be entitled.

1. Name of Participant ("Study Subject")

____

2. Title of Research Study

The impact of acute psychological stress on cognitive functioning

3. Principal Investigator and Telephone Number(s)

Kevin G. F. Thomas, Ph.D.
Department of Psychology
University of Cape Town
021-650-4608

Robyn Human, B.Soc.Sc.
Honours Candidate
Department of Psychology
University of Cape Town
021-788-5536

Michelle Henry, B.Sc.
Honours Candidate
Department of Psychology
University of Cape Town
021-551-6534

Health Sciences Faculty
Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Tel: 021-406-6338

Fax: 021-406-6411
Email: lamees.emjedi@uct.ac.za

4. What is the purpose of this research study?

The purpose of this research study is to better understand how exposure to acute psychological stress affects cognitive functioning. More specifically, we are interested in individual differences in cognitive responses to acute psychological stress.

5. What will be done if you take part in this research study?

This study requires you to be take part in two research sessions on two consecutive days. On the first day you will be required to complete a number of memory-based tasks. On the second day you may be required to complete a 20-minute presentation which will be followed by another series of memory based tasks. Throughout the study your levels of stress will be assessed through the collection of heart rate measurements and saliva samples with the aid of a cotton swab. These saliva samples will be used to analyse levels of salivary cortisol.

6. What are the possible discomforts and risks?

If you are one of the participants selected to complete the 20-minute presentation, you may be placed in a mildly stressful situation involving public speaking. There are no other discomforts and risks associated with participation in the study.

7. What are the possible benefits of this study?

One major benefit of this study is that scientists, and society in general, will have better understanding of the effects of acute psychological stress on cognitive functioning. This knowledge can then be applied to many different individuals and situations, including students who are taking exams, business managers who have to present to their boards, and so on.

8. Can you withdraw from this research study and if you withdraw, can information about you still be used and/or collected?

You may withdraw your consent and stop participation in this study at any time. Information already collected may be used.

9. Once personal information is collected, how will it be kept confidential in order to protect your privacy and what protected health information about you may be collected, used and shared with others?

Information collected will be stored in locked filing cabinets or in computers with security passwords. Only certain people - the researchers for this study and certain University of Cape Town officials - have the legal right to review these research records. Your research records will not be released without your permission unless required by law or a court order.

If you agree to be in this research study, it is possible that some of the information collected might be copied into a "limited data set" to be used for other research purposes. If so, the limited data set may only include information that does not directly identify you.

10. Signatures

As a representative of this study, I have explained to the participant the purpose, the procedures, the possible benefits, and the risks of this research study; the alternatives to being in the study; and how the participant's protected health information will be collected, used, and shared with others:

____ _____________________
Signature of Person Obtaining Consent and Authorization Date

You have been informed about this study's purpose, procedures, and risks; how your protected health information will be collected, used and shared with others. You have received a copy of this form. You have been given the opportunity to ask questions before you sign, and you have been told that you can ask other questions at any time.

You voluntarily agree to participate in this study. You hereby authorize the collection, use and sharing of your protected health information. By signing this form, you are not waiving any of your legal rights.

____ _____________________
Signature of Person Consenting and Authorizing Date

Please indicate below if you would like to be notified of future research projects conducted by our research group:

______________ (initial) Yes, I would like to be added to your research participation pool and be notified of research projects in which I might participate in the future.

Method of contact:
Phone number: _________________________________
E-mail address: _________________________________
Mailing address: _________________________________
____
____

FOOTNOTES

[1]Women in the late luteal phase have comparable saliva cortisol stress responses to men, whereas women in the follicular phase of the menstrual cycle, or those taking oral contraceptives, have significantly lower free cortisol responses (Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999). For the above reasons it was vital that women in this study were in the late luteal phase during the experimental procedure.

[2]Several studies have shown that time of day is a crucial factor when performing experiments featuring cortisol measurements. Evidence shows that HPA axis responses depend on the time of day, with larger cortisol responses in the afternoon and evening compared to the morning (Kudielka, Schommer, Hellhammer, & Kirschbaum, 2004). HPA axis activity follows a circadian rhythm, with highest hormone levels in the early morning hours followed by continually decreases over the course of the day (Kirschbaum & Hellhammer, 1994). These high levels of cortisol in the morning result in smaller endocrine responses to pharmacological or environmental provocations. Furthermore, studies have shown that mood changes may be more pronounced in the morning compared to the afternoon and evening, with a more positive mood in the morning (Kudielka et al., 2004). This may in part explain the why cortisol levels are elevated during the morning.

[3]These factors may cause fluctuations in baseline cortisol levels (Kirschbaum et al., 1993), which need to be kept constant prior to experimentation.

Table 1.

Measures of Stress

	Group	
	Stress *n* = 6	Control *n* = 12
STAI Trait	37.5 (11.34)	37.92 (8.95)
STAI State		
Baseline	36.17 (5.53)	34.00 (10.27)
Post-manipulation	38.50 (9.00)	27.50 (6.68)
End	30.50 (4.09)	26.42 (5.38)
Cortisol		
Baseline	2.27 (2.30)	3.39 (4.09)
Post-manipulation	4.91 (4.15)	1.66 (1.68)
End	1.15 (0.92)	1.87 (3.24)

Note. Means are presented with standard deviations in parentheses.

Self-Reported Measures of Anxiety

Table 2.

Recognition of Each Item Type as a Function of Test Type

	Group			
	Stress		Control	
	n = 6		*n* = 12	
	Day 1	Day 2	Day 1	Day 2
Standard Test				
Both hits	0.87 (0.11)	0.82 (0.10)	0.92 (0.06)	0.88 (0.09)
Red word hits	0.65 (0.19)	0.67 (0.12)	0.67 (0.14)	0.61 (0.19)
Picture hits	0.74 (0.15)	0.81 (0.10)	0.76 (0.17)	0.78 (0.12)
New FAs	0.24 (0.25)	0.37 (0.20)	0.14 (0.10)	0.40 (0.19)
Picture Test				
Both hits	0.65 (0.27)	0.61 (0.29)	0.79 (0.14)	0.74 (0.13)
Red word FAs	0.21 (0.19)	0.27 (0.14)	0.18 (0.20)	0.20 (0.18)
Picture hits	0.75 (0.21)	0.65 (0.31)	0.73 (0.18)	0.68 (0.18)
New FAs	0.22 (0.32)	0.28 (0.25)	0.11 (0.13)	0.21 (0.18)
Red Word Test				
Both hits	0.60 (0.19)	0.49 (0.21)	0.60 (0.25)	0.59 (0.21)
Red word hits	0.76 (0.19)	0.65 (0.13)	0.72 (0.15)	0.65 (0.22)
Picture FAs	0.37 (0.21)	0.35 (0.15)	0.36 (0.18)	0.47 (0.17)
New FAs	0.44 (0.34)	0.53 (0.15)	0.36 (0.36)	0.45 (0.25)

Note. Standard errors of each mean are in parentheses.; Both hits = correct responses to items presented as both pictures and words during the study phase; Red Word hits = correct responses to items presented as only red words during the study phase; Picture hits = correct responses to items presented as only pictures during the study phase; New FAs = false alarms to items that were never presented during the study phase; Red Word FAs = false alarms in the Picture test whereby participants incorrectly labelled an item that was studied as a Red Word during the study phase as being studies as a Picture; Picture FAs = false alarms in the Red Word test whereby participants incorrectly labelled an item that was studied as a Picture during the study phase as being studied as a Red Word.

Table 3.

Combined Scores of Stress and Control Groups Recognition of Each Item Type as a Function of Test Type for Day 1 Analyses

	n = 18
Standard Test	
Both hits	0.91 (0.08)
Red word hits	0.66 (0.15)
Picture hits	0.75 (0.16)
New FAs	0.17 (0.16)
Picture Test	
Both hits	0.74 (0.20)
Red word FAs	0.19 (0.19)
Picture hits	0.74 (0.19)
New FAs	0.15 (0.21)
Red Word Test	
Both hits	0.60 (0.23)
Red word hits	0.73 (0.16)
Picture FAs	0.36 (0.18)
New FAs	0.39 (0.34)

Note. Standard errors of each mean are in parentheses.; Both hits = correct responses to items presented as both pictures and words during the study phase; Red Word hits = correct responses to items presented as only red words during the study phase; Picture hits = correct responses to items presented as only pictures during the study phase; New FAs = false alarms to items that were never presented during the study phase; Red Word FAs = false alarms in the Picture test whereby participants incorrectly labelled an item that was studied as a Red Word during the study phase as being studies as a Picture; Picture FAs = false alarms in the Red Word test whereby participants incorrectly labelled an item that was studied as a Picture during the study phase as being studied as a Red Word.

Table 4.
True Memory Performance on Criterial Recollection Tests

	t	*df*	*p*	Cohen's *d*
Picture Recollection Test				
Comparison:				
Both hits vs Picture hits	0.07	34	0.472	0.03
Red Word Recollection Test				
Comparison:				
Both hits vs Red Word hits	-2.05	30.6	0.024	0.68
Picture hits (Picture Test vs Red Word hits (Red Word Test)	0.12	34	0.452	0.04

Note. All tests were one-tailed.

Table 5

False Memory Performance on Criterial Recollection Tests

	t	*df*	*p*	Cohen's *d*
Picture Recollection Test				
Comparison:				
Red Word FAs vs New FAs	0.66	34	0.257	0.22
Red Word Recollection Test				
Comparison:				
Picture FAs vs New FAs	0.33	25.9	0.373	0.11
New FAs (Picture Test) vs New FAs (Red Word Test)	-2.60	28.18	0.007	0.87
Red Word FAs (Picture Test) vs Picture FAs (Red Word Test)	-2.75	34	0.005	0.92

Note. All tests were one-tailed.

Table 6.
Comparing Picture and Word hits within the Stress Group

	t	*df*	*p*	Cohen's *d*
Standard Recognition Test				
Comparison:				
Picture hits vs Red Word hits	2.15	10	0.029	1.27
Picture hits (Picture test) vs Red Word hits (Red Word test)	-0.05	10	0.480	0.00

Note. All tests were one-tailed.

Table 7.
Comparing Picture and Word hits within the Control Group

	t	*df*	*p*	Cohen's *d*
Standard Recognition Test				
Comparison:				
Picture hits vs Red Word hits	2.61	22	0.008	1.07
Picture hits (Picture test) vs Red Word hits (Red Word test)	0.38	22	0.353	0.15

Note. All tests were one-tailed.

Table 8

Between Group Comparisons of False Memory Performance on Recognition Tests

	t	*df*	*p*	Cohen's *d*
New FAs				
Comparison:				
Standard Recognition Test	-0.33	16	0.374	0.16
Picture Criterial Recollection Test	0.64	16	0.265	0.32
Red Word Criterial Recollection Test	0.77	16	0.226	0.37
Picture FAs (Red Word Test)	-1.48	16	0.08	0.73
Red Word FAs (Picture Test)	0.88	16	0.195	0.48

Note. All tests were one-tailed.

Table 9
Within Group Comparisons of False Memory Performance on Criterial Recollection Tests

	t	*df*	*p*	Cohen's *d*
Stress Group:				
Picture FAs (Red Word Test) vs Red Word FAs (Picture Test)	1.00	10	0.170	0.55
Control Group:				
Picture FAs (Red Word Test) vs Red Word FAs (Picture Test)	3.94	22	*0.001	1.54

Note. All tests were one-tailed.
p < .05, two-tailed. **p < 001, one-tailed.

Table 10

Stress Groups True and False Memory Performance on Recognition Tests

	t	*df*	*p*	Cohen's *d*
Standard Test				
Comparison:				
Both Hits	-0.84	5	0.219	0.48
Red Word Hits	0.53	5	0.311	0.13
Picture Hits	1.65	5	0.08	0.55
New FAs	-1.86	5	0.122	0.57
Red Word Test				
Comparison:				
Both Hits	-1.29	5	0.127	0.55
Red Word Hits	-1.21	5	0.140	0.68
Picture FAs	-0.21	5	0.844	0.11
New FAs	0.55	5	0.605	0.34
Picture Test				
Comparison:				
Both Hits	-1.94	5	0.06	0.14
Red Word FAs	1.63	5	0.164	0.36
Picture Hits	1.18	5	0.146	0.38
New FAs	-1.30	5	0.251	0.21

Note. Test comparing hits were one-tailed; tests comparing false alarm scores were two-tailed.

Table 11

Control Groups True and False Memory Performance on Recognition Tests

	t	*df*	*p*	Cohen's *d*
Standard Test				
Comparison:				
Both Hits	2.12	12	0.029	0.52
Red Word Hits	-1.21	12	0.127	0.36
Picture Hits	0.35	12	0.365	0.14
New FAs	-5.38	12	*0.0002	0.92
Red Word Test				
Comparison:				
Both Hits	0.18	12	0.430	0.04
Red Word Hits	0.83	12	0.212	0.37
Picture FAs	2.19	12	0.051	0.63
New FAs	-0.74	12	0.477	0.29
Picture Test				
Comparison:				
Both Hits	-1.19	12	0.129	0.37
Red Word FAs	0.34	12	0.740	0.11
Picture Hits	-1.21	12	0.126	0.28
New FAs	1.92	12	0.081	0.64

Note. Test comparing hits were one-tailed; tests comparing false alarm scores were two-tailed.
**p < 001, one-tailed.

Table 12.

Difference Scores (Day 1- Day2) for Each Item Type as a Function of Test Type

	Group	
	Stress	Control
	n = 6	*n* = 12
	Difference Score	Difference Score
Standard Test		
Both hits	0.05 (0.14)	0.05 (0.08)
Red word hits	-0.03 (0.13)	0.05 (0.15)
Picture hits	-0.08 (0.11)	-0.02 (0.17)
New FAs	-0.13 (0.17)	-0.26 (0.17)
Picture Test		
Both hits	0.04 (0.05)	0.05 (0.14)
Red word FAs	-0.06 (0.09)	-0.02 (0.17)
Picture hits	0.10 (0.22)	0.05 (0.14)
New FAs	-0.06 (0.10)	-0.10 (0.19)
Red Word Test		
Both hits	0.10 (0.20)	0.01 (0.14)
Red word hits	0.11 (0.22)	0.06 (0.26)
Picture FAs	0.01 (0.16)	-0.11 (0.18)
New FAs	-0.09 (0.40)	-0.08 (0.39)

Note. Standard errors of each mean are in parentheses.; Negative difference scores denote an increase in memory from Day 1 to Day 2; positive difference scores denote a decrease in memory from Day 1 to Day 2.

Table 13.

Comparison of True Memory Difference Scores between the Stress and Control Group

	t	*df*	*p*	Cohen's *d*
Standard Test				
Comparison:				
Both Hits	-0.003	16	0.499	0.00
Red Word Hits	-1.11	16	0.141	0.55
Picture Hits	-0.77	16	0.227	0.39
Red Word Test				
Comparison:				
Both Hits	1.24	16	0.117	0.62
Red Word Hits	0.39	16	0.351	0.19
Picture Test				
Comparison:				
Both Hits	-0.15	15.38	0.443	0.07
Picture Hits	0.66	16	0.259	0.33

Note. All tests were one-tailed.

.